Connectivity between MPAs: selecting appropriate taxa and assessing genetic connectivity in two benthic marine invertebrates

Submitted by Thomas Lowther Jenkins to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Biological Sciences, October 2018.

(Signature) .....

Abstract

Connectivity is fundamental for the persistence of many populations of marine species and is formally identified as one of five key principles for designing an ecologically coherent network of Marine Protected Areas (MPAs) in European waters. However, the process of assessing connectivity between MPAs, and which taxa to include in assessments of connectivity, is challenging. Managers of MPAs have typically concentrated their efforts on species that are endangered or rare, or on so-called 'umbrella', 'keystone' or 'flagship' species; however, these species may not always be the best candidates for assessing connectivity of a MPA network. In this thesis, a meta-analysis was firstly conducted to study genetic patterns across a broad range of coastal marine taxa in the northeast Atlantic. This meta-analysis provided insights into the biological and methodological information needed to ascertain which taxa may be considered as good candidates for assessing genetic connectivity between MPAs across Britain and the wider northeast Atlantic. The knowledge gained from this literature survey facilitated the design of a set of criteria that identified ideal traits of a candidate species for assessments of genetic connectivity between MPAs; subsequently, based on these criteria, two species were selected to assess connectivity between MPAs in the British network: the pink sea fan (*Eunicella verrucosa*) and the European lobster (*Homarus gammarus*). Using 13 microsatellites and 3,743 SNPs, the results for the pink sea fan indicated the presence of three distinct genetic groups, partitioned between sites from western Ireland, southern Portugal and Britain-France. For the European lobster, 86 SNPs indicated strong genetic differentiation between the northeast Atlantic, the middle Mediterranean and the eastern Mediterranean (Aegean Sea). In addition, there was a pronounced genetic cline across the northeast Atlantic, suggesting that connectivity in the European lobster follows a stepping-stone model of dispersal, which was supported by simulations of larval dispersal. Taken together, the results from these two studies suggests that the MPA network in Britain is sufficient to maintain connectivity in the pink sea fan and the European lobster, and possibly other species living in comparable habitats with similar life histories and dispersal traits. Moreover, the criteria applied in this thesis to select species appears to facilitate the identification of ideal surrogate taxa to assess connectivity between MPAs, which could easily be applied to assessments of MPA network connectivity in other seas and oceans around the world.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I know it is a cliche in these acknowledgement sections, but there are so many people who have helped to make this thesis a reality. Collaboration is, to me, one of the main traits of successful science, and this has been central to each component of research carried out in this thesis. This is particularly relevant to all of the people from across the UK and Europe who kindly gave up their time to collect tissue samples of lobster for me, and for my colleagues and friends who trained me in wet-lab molecular biology techniques.

Before I dive into the plethora of acknowledgements, I'd like to start by thanking the primary funders of my PhD, the NERC GW4+ DTP, the University of Exeter and Natural England (in particular Emma Dade and James Highfield), for which this PhD would not have been possible without your support. In addition, a special thank you to the Genetics Society and the British Ecological Society who have generously supported my attendence to several conferences and training workshops, and extra fieldwork opportunities, which has enhanced both my PhD and my personal development as an independent researcher.

Secondly, I would like to thank my PhD supervisor, Jamie Stevens, for initially entrusting me with this PhD! And for being an incredibly dedicated and motivating supervisor; our insightful discussions have always made me think critically about the science. Also, thank you to all of the members of the Molecular Ecology and Evolution Group (MEEG) for your ideas and advice in and out of the lab. In particular, I would like to thank Josie Paris for your help and support in bioinformatics and Guy Freeman for the exceptional species illustrations used in some of the figures in this thesis. In addition, a massive thanks to Lauren Ames and Graham Thomas for passing on your wet-lab tips and tricks – they won't be forgotten! And a thank you to Rita Castilho, whose knowledge and expertise was extremely valuable for analysing, interpreting and publishing data from chapter two.

Thirdly, and most importantly for the lobster project, a huge thank you to all the people who kindly collected or provided me with lobster tissue samples for genetic analysis: James Wood (Bridlington), Jonas Seafood Ltd (Cromer), D.R.Collin & Sons (Eyemouth), Donald MacLennan (Outer Hebrides), Jack Emmerson (Isle of Man), W Harvey & Sons (Isles of Scilly), Sion Williams (Llyn Peninsula), Dawson Shearer and Matt Coleman (Orkney), Charlie Ellis (Padstow), DASH Shellfish (Pembrokeshire), Mark Hamilton (Shetland), Monteum Ltd (Shoreham-by-sea), Douglas Craigie (Sula Sgeir), Peter Gay and Don Thompson (Channel Islands), Fabienne Legall (Île de Ré), Roland Krone (Helgoland), Owen Doyle (Arranmore Island and Donegal Bay), Ron Randalls (Cork Harbour), John Hickey (Hook

Peninsula), Majbritt Bolton-Warberg (Kilkieran Bay), Declan Nee (Mullet Peninsula), Vera O'Donovan (Ventry), Alexandros Triantafyllidis (Greece), Sara Saba and Ulisse Murru (Sardinia), Tonje Knutsen (Flodevigen and Tvedestrand), Carl Andre (Gullmarfjord, Kavra and Singlefjord), Ellika Faust (Lysekill) and Gonzalo Perez Benavente (Vigo). This study would not have been possible without you. I would like to especially thank the team at the National Lobster Hatchery for your support and enthusiasm, particularly Charlie Ellis for helping me collect lobster samples and for providing insightful ideas and comments for chapter six.

Finally, I never would have got this far if it wasn't for my friends, girlfriend, family, coffee and Marvel films. First, a massive shout out to past and present members of coffee club – relinquishing my admin duties and deciding a successor will be difficult, but I trust the true heir will be revealed at some point. Also, a shout out to my friend and fellow PhD submitter, Matt Jones, whose competitive FIFA duals and banter have kept me sane. Second, to my girlfriend, Abbi, who has continued to stick with me over the last few years, and particularly the last several months – coming home everyday has put a smile on my face. Lastly, to my amazing family, Mum, Dad and Bubus, who have supported me throughout this perpetual student phase, even when you've had no idea what I'm talking about! Spending weekends at home and Facetiming in the evenings has kept me going and I can't thank you enough for your support.

Abbreviations

AU	Adaptive unit
BAP	Biodiversity Action Plan
BIM	Bord Lascaigh Mhara
bp	Base pair
CBD	Convention on Biological Diversity
CU	Conservation unit
DAPC	Discriminant Analysis of Principal Components
DEFRA	Department for Environment Food and Rural Affairs
DNA	Deoxyribonucleic acid
EEZ	Exclusive economic zone
ESU	Evolutionary significant unit
HWE	Hardy-Weinberg Equilibrium
IBD	Isolation-by-distance
IFC	Integrated Fluidic Circuit
IFCA	Inshore Fisheries and Conservation Authorities
IUCN	International Union for Conservation of Nature
IUU	Illegal unreported and regulated fishing
JNCC	Joint Nature Conservation Committee
Ka	Thousand years ago
LD	Linkage disequilibrium
LFU	Lobster Fishery Unit
LGM	Last Glacial Maximum
Ma	Million years ago
MCMC	Markov chain Monte Carlo
MCZ	Marine Conservation Zone
MPA	Marine Protected Area
MSFD	Marine Strategy Framework Directive
mtDNA	Mitochondrial DNA
MU	Management unit
NCMPA	Nature Conservation Marine Protected Area
NE	Natural England
NERC	Natural Environment Research Council
NTZ	No Take Zone
OSPAR	Oslo/Paris Convention
PCR	Polymerase chain reaction
PLD	Pelagic larval duration

RADseq	Restriction-site associated DNA sequencing
RNA	Ribonucleic acid
SAC	Special Areas of Conservation
SNP	Single nucleotide polymorphism
SPA	Special Protected Area
UK	United Kingdom

Chapter 1: General introduction

Understanding the biology, ecology and evolution of marine fauna and flora has never been more readily achievable. Previously, scientific knowledge of marine organisms routinely trailed their terrestrial counterparts, which likely arose from the obvious obstacles associated with observing, sampling and studying organisms in marine systems. However, decades of technological advances now permit the exploration of many marine and estuarine environments, substantially narrowing the knowledge gap between marine and terrestrial systems.

The study of connectivity is one such discipline that has benefited from these technological advances. Connectivity research links a number of different fields in ecology and evolution including animal behaviour, population dynamics, genetic structure analysis and adaptation to local environments (Kool et al. 2013). The discipline has also emerged as a key component of applied science and conservation; for example, understanding connectivity has been vital for designing networks of protected areas (Jones et al. 2007), tracking pathways of invasive species (Pérez-Portela et al. 2012), managing fisheries resources (Fogarty & Botsford 2007; Kough et al. 2013), and monitoring the effects of climate change (Gerber et al. 2014).

This thesis takes advantage of the recent developments in both marine connectivity and DNA sequence technology to investigate patterns of genetic diversity and connectivity in two coastal marine invertebrates. Although this PhD ultimately aims to generate novel population genetic data for the pink sea fan (*Eunicella verrucosa*) and the European lobster (*Homarus gammarus*), the overarching theme of this thesis is steered towards informing marine conservation and management. Therefore, as well as exploring the population biology and ecology of these two species, a central premise is to integrate and translate the findings into usable forms of evidence that can inform practitioners in the fields of marine protected area designation and implementation and fisheries management.

1.1 Marine connectivity

Marine connectivity is the study of dispersal and immigration between populations of marine organisms. In particular, it has been defined as the extent to which populations inhabiting different parts of a species' range are linked by the movement of eggs, propagules, larvae, juveniles or adults (Palumbi 2003). High

connectivity between two populations implies that many individuals are exchanged between these populations, whereas limited or no exchanges implies restricted or no connectivity. This definition, however, does not consider that an individual may disperse to a new population but die upon (or shortly after) reaching its destination. Realised connectivity (also used interchangeably with functional or effective connectivity) is where an individual successfully navigates across habitat patches and concludes in successful settlement and reproduction in the new population (Watson et al. 2010; Kool et al. 2013). Lowe & Allendorf (2010) also distinguished demographic and genetic connectivity (section 1.1.4), which essentially identifies and explores the influence of, and links between, short temporal scale population dynamics and large-scale evolutionary processes (Hidalgo et al. 2017).

Marine connectivity is fundamental for the persistence of many populations of marine species. It is particularly important for sessile and non-motile marine fauna that rely primarily on ocean currents to disperse to nearby or distant populations (Cowen & Sponaugle 2009). This ability to travel potentially vast distances via ocean currents likely explains the large ranges associated with many marine fauna; it also increases the capacity for species to expand their range by colonising new habitat patches that become available. Indeed, connectivity can be extremely important for the long-term stability of meta-populations and for recolonising habitats that have experienced local extirpation (Hanski 1998). Hence, one of the aims of marine conservation is to establish marine reserves of adequate size and spacing to help maintain natural connectivity by reducing human disturbance in key or vulnerable areas (Fogarty & Botsford 2007; Botsford et al. 2009).

Populations persist when self-recruitment and immigration equal or exceed mortality and emigration (Cowen & Sponaugle 2009). Populations are sinks when the net import of individuals is greater than the net export of individuals; conversely, source populations export more individuals than they receive. Sinks that have low local recruitment (i.e. low survival and birth rates of residents) are assumed to benefit considerably from immigrant subsidy to maintain stable populations (Runge et al. 2006). This is unlikely the case for pseudo-sinks, however, which can be independently viable populations, but appear as 'true' sinks because asymmetrical immigration can depress fecundity or increase mortality because of density-dependent processes (Watkinson & Sutherland 1995). Yet, distinguishing pseudo-sink populations from genuine sinks is seldom straightforward. Nevertheless, identifying source and sink populations can be very important for designing networks of protected areas (Jones et al. 2007) and for the spatial management of marine fisheries (Fogarty & Botsford 2007).

1.1.1 Models of connectivity

Simple but logical models attempt to conceptualise connectivity (Fig. 1), forming a theoretical baseline for investigating connectivity in natural populations. The three most widely considered models are the island model (Fig. 1A; Wright 1932), the hierarchical island model (Fig. 1B; Slatkin & Voelm 1991) and the stepping-stone model (Fig. 1B; Kimura & Weiss 1964). In the island model, all populations are connected via a constant migration rate, whereas in the hierarchical island model, a group of populations exchange migrants with each other at a much higher rate compared with other populations or groups of populations (Slatkin & Voelm 1991). In other words, in the hierarchical island model, connectivity within a subset of populations is higher compared with some other external populations. In comparison to these two models, the stepping-stone model considers the spatial arrangement of populations and proposes that the migration rate is stronger between populations that are spatially closer together. In effect, the model suggests that in each generation, an individual can migrate one 'step' in any direction to an adjacent population, and then the process repeats itself in subsequent generations (Kimura & Weiss 1964). This model is somewhat realistic for species with large ranges and is usually associated with isolation-by-distance (IBD), a term coined by Wright in the 1940s to describe the distribution of genetic variation over a geographic region (Wright 1943).

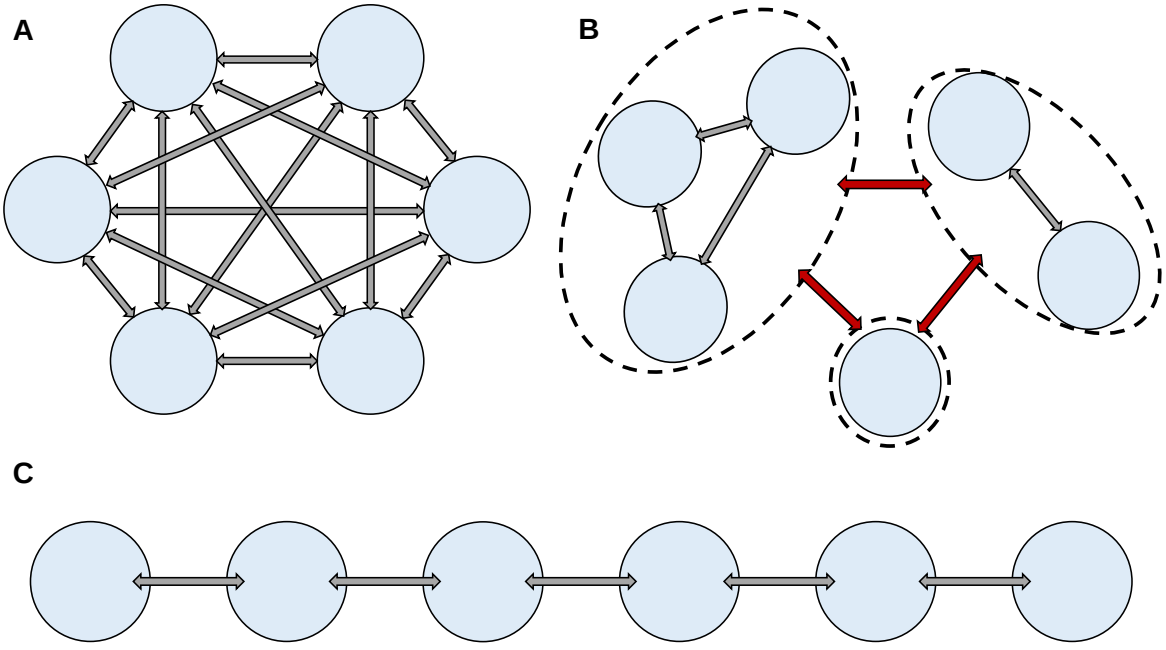


Figure 1: Three models of connectivity: (A) the island model, (B) the hierarchical island model and (C) the stepping-stone model. Blue circles represent hypothetical populations. Arrows denote migration rates; red arrows represent much less migration than grey arrows.

For marine species with pelagic larval phases, the island model in essence describes a situation in which larvae are drawn from a well-mixed pool of larvae that have originated from multiple populations across the species geographical range. This is likely due to high dispersal capacity of the larvae, coupled with adequate oceanography to facilitate mixing, which ultimately results in high connectivity over large spatial scales. In comparison, the stepping-stone model describes a scenario in which the larval dispersal capacity is insufficient to travel from one habitat patch to a distant habitat patch in a single dispersal event. Instead, larvae more often disperse to neighbouring populations which may lead to the development of IBD patterns, whereby connectivity is stronger between populations that are closer together than between populations that are further apart, resulting in higher genetic similarity between populations that are spatially closer to one another.

1.1.2 Drivers of connectivity

The marine environment presents many opportunities for dispersal within and between populations. Benthic marine organisms are typically sedentary or immobile as adults and, therefore, rely on the natural oceanography around them to disperse their eggs or larvae (Cowen & Sponaugle 2009). During this pelagic larval phase, the larvae form a temporary local population in the plankton, and then drop out of the water column after a period of drifting to find suitable habitat for settlement. In the early days, this high potential for long-distance dispersal led to the belief that marine populations were universally well-connected and 'open' over ecological time scales (Cowen et al. 2000). This was initially supported by a number of genetic studies that reported large homogenous populations over regional to basin-wide spatial scales (Cowen et al. 2007, and references therein). However, recent research on many different marine taxa, and reconsideration of previous evidence using more modern genetic markers, suggests that limited dispersal and connectivity may be more prevalent in marine environments than previously thought (Hauser & Carvalho 2008; Hellberg 2009), challenging the paradigm that populations of marine species are universally open over ecological time scales.

The bipartite life cycle of many benthic marine organisms means that the total distance travelled during dispersal is primarily driven by environmental and ontogenic factors during the pelagic larval phase, which is highly variable across taxa and can last anywhere from hours to months (Shanks 2009). As eggs or larvae are expelled into the water column above the seafloor, they are immediately

subjected to the velocity and direction of the ambient ocean currents. Larvae that have limited ability to swim against water masses are completely dependent on the ocean currents to disperse away from their natal origin. Logically, this suggests that interspecific larvae released from the same habitat with a similar pelagic larval duration (PLD) may be expected to have equivalent dispersal potential. However, because of the diverse life history strategies employed by marine organisms (section 1.1.3), and the influence of biotic (e.g. predation, spawning timing) and abiotic factors (e.g. habitat availability), dispersal distance and connectivity can vary substantially between species (Bradbury et al. 2008; Cowen & Sponaugle 2009). For example, nutrient availability and sea temperatures play a critical role in dictating the PLD of some species (e.g. McCormick & Moloney 1995), which could also mean that projected increases in sea temperature with global climate change may have tangible effects on future patterns of marine connectivity (O'Connor et al. 2007).

A break in connectivity can occur when certain conditions restrict dispersal between populations. For example, glaciers during the Pleistocene epoch isolated many terrestrial populations of northwest American taxa, preventing movement between populations (Shafer et al. 2010). In the marine environment, the formation of a land barrier would instantly prevent movement between populations either side of this barrier; such an example exists in Central America, the Isthmus of Panama (Fig. 2), which separates the Pacific and Atlantic oceans. The Isthmus of Panama formed around 2.8 million years ago (Ma) during the Cenozoic era (O'Dea et al. 2016). This introduced a land barrier to marine species but interestingly removed a marine barrier to terrestrial species from North and South America (Fig. 2); this meant that migration and gene flow was nullified in the marine realm but became possible in the terrestrial realm. For marine species, this vicariant event resulted in reproductive isolation and independent evolution between populations of the two oceans, leading to gradual divergence over time and the beginnings of allopatric speciation in some taxa (Lessios 2008).

More subtle barriers can also impede dispersal in the marine environment, such as eddies, fronts, deep-water and environmental gradients (e.g. salinity and temperature). The magnitude of the effect of these barriers to dispersal and connectivity can vary depending on a species' life history and the strength of the barrier. For example, the Almeria-Oran front in the Mediterranean Sea has been found to constrain connectivity in some, but not all, species (Patarnello et al. 2007; Pascual et al. 2017). Moreover, deep-water has been suggested to hinder connectivity between the central and eastern Pacific in sedentary shallow water species whose larvae cannot stay in the plankton long enough to bridge the stretch

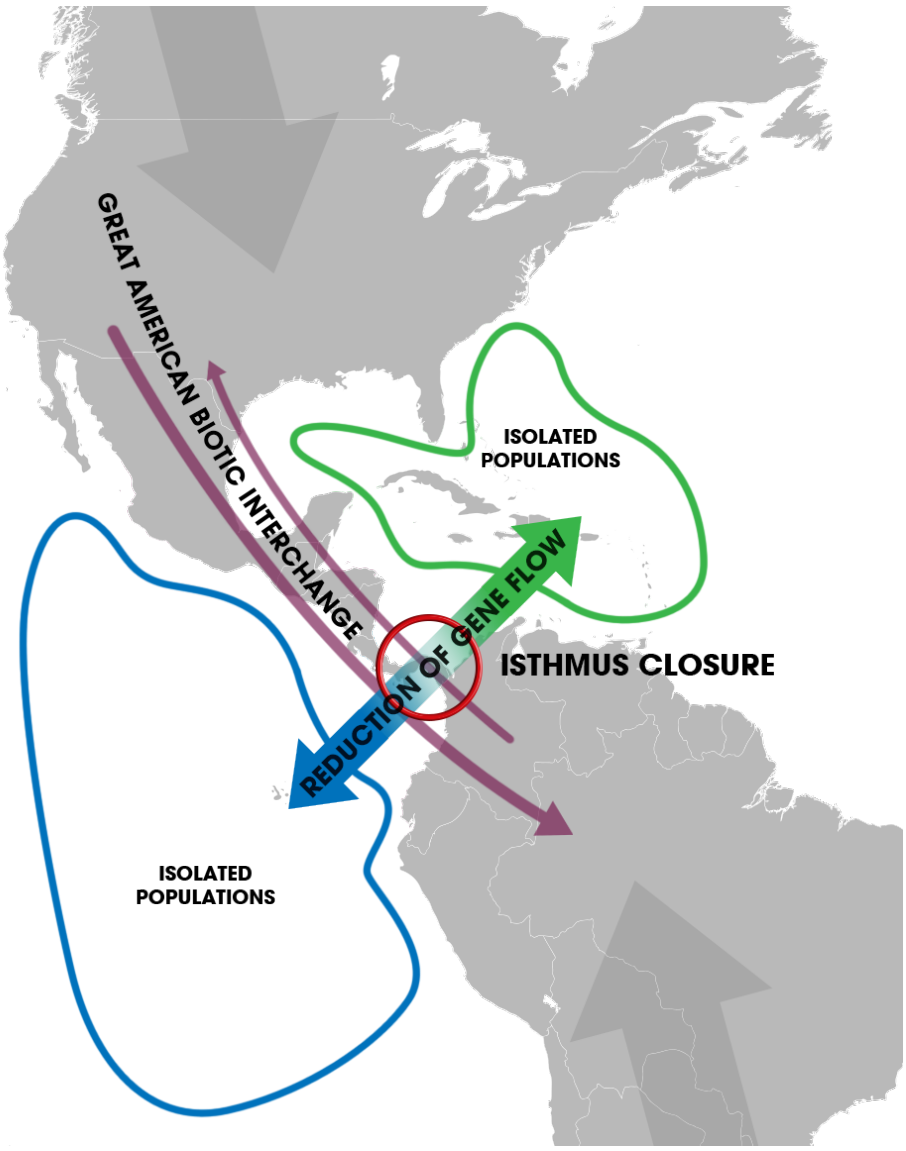


Figure 2: The Isthmus of Panama. The formation of the Isthmus of Panama introduced a land barrier to populations of marine species but removed a marine barrier to terrestrial species from North and South America. Image created by Andrew Z. Colvin under a CC BY-SA 4.0 licence.

of deep-water (Lessios 2012). On the other hand, the circular currents of eddies may promote self-recruitment and connectivity between habitats located within the eddy system (Sponaugle et al. 2005).

Lastly, there are mounting reports of invasive species that have likely arisen from unnatural dispersal, which is possibly a consequence of intentional (e.g. aquaculture) or unintentional (e.g. lionfish in the Caribbean via ballast water) human-mediated translocations (Lowry et al. 2013). Unfortunately, invasive species can have negative impacts on biodiversity, ecosystem functioning and ecosystem services (e.g. lionfish, Green et al. 2012), which has led to the implementation of

national and international programs to mitigate current and future impacts (Pysek & Richardson 2010). From a conservation genetics perspective, human-mediated translocations can also make inferences about population structure difficult because it can produce abnormal phylogeographic patterns that no longer reflect natural processes (Ni et al. 2012). In addition, continuing with the theme of human-mediated dispersal, there has also been an increase in 'ocean sprawl' over the last few decades (Firth et al. 2016); ocean sprawl is the proliferation of artificial structures in the sea. These structures, including offshore marine renewable energy devices and wrecks, have been suggested to provide artificial stepping stones for marine fauna (Krone & Schröder 2011; Adams et al. 2014); however, this scheme is still in its infancy and more research is needed to quantify the effects (positive or negative) of ocean sprawl on ecological connectivity in the marine environment (Bishop et al. 2017). In summary, when studying benthic marine organisms, it is also important to consider potential routes of anthropogenic dispersal and its possible impacts when exploring patterns of dispersal and connectivity.

1.1.3 Dispersal, recruitment and life history strategies

Dispersal is defined by Cayuela et al. (2018) as the movement of individuals from their natal patch of birth to their first breeding site. Effective dispersal is similarly defined, except that the disperser spends enough time at the new site to successfully reproduce and transmit its genes to the next generation. In population genetic studies, "dispersal" and "migration" are often considered synonyms; however, Cayuela et al. (2018) highlight that these are two distinct concepts, in which dispersal (effective or non-effective) is usually a one-way process, whereas migration implicates recurrent, two-way movements.

Dispersal success can vary considerably across space and time as the marine environment is not uniform, but a highly dynamic heterogenous oceanic landscape (i.e. seascape) which presents challenges for dispersing individuals. Benthic marine organisms have evolved a diverse array of life history strategies and behaviours (e.g. diel vertical migration) to deal with these challenges (Levin 2006), which interact with oceanography to ultimately define dispersal success (Cowen & Sponaugle 2009). Some modes of reproduction and development are, theoretically, expected to hamper or increase larval dispersal (Table 1); the most common of these for benthic marine species are briefly discussed.

Direct developing organisms lack a pelagic larval phase, with offspring emerging from the parent as larvae and then settling after expulsion, or offspring

Table 1: Common life history strategies of benthic marine species and their expected effects on dispersal potential.

Strategy	Description	Dispersal potential
Development		
Direct	Develop inside parents or from eggs. Larvae may look like adult form straight after hatching.	Low
Brooder	Eggs fertilised inside maternal adult. Larvae ejected and settle nearby.	Low
Surface brooder	Eggs fertilised inside maternal adult. Larvae ejected and are phototaxic.	Low
Broadcast spawner	Discharges gametes into the water column.	High
Pelagic phase		
No PLD	No pelagic larval duration.	Low
Lecithotrophic	Disperse via ocean currents, provided a yolk sac for nutrition.	Medium
Planktotrophic	Disperse via ocean currents, feed while in the plankton.	High

are devoid of a larval stage altogether and emerge as a miniature version of the adults. For example, brooders release sperm (assumed to be negatively buoyant) which fertilise eggs inside the maternal adult and larvae are then ejected and settle nearby; this is the case for some sexually reproducing corals (e.g. *Corallium rubrum*, Ledoux et al. 2010). A similar strategy, surface brooding, is where larvae show phototaxis behaviour (attraction or repulsion to light), whereby larvae swim towards or away from light at the ocean surface; such a strategy is adopted by larvae of the red gorgonian (*Paramuricea clavata*), which display negative phototaxis and remain in suspension for only a few minutes before descending to the seafloor to settle (Mokhtar-Jamai et al. 2011). A slightly different method of direct development occurs in oviparous skates and rays, in which the eggs are internally fertilised and egg cases are deposited on the seafloor where, after a period of incubation, fully developed young emerge (e.g. *Raja clavata*, Chevolot et al. 2006). In comparison to these direct strategies, broadcast spawners synchronously discharge their gametes into the water column, which are externally fertilised and then swept away by the surrounding currents (e.g. the octocoral *Eunicella verrucosa*, Munro 2004). Because of the duration spent drifting in ocean currents, broadcast spawning is expected to have much higher dispersal potential than direct modes of development (Table 1). During the pelagic phase, the meroplanktonic larvae are either lecithotrophic (provided with a yolk sac) or planktotrophic (feed in the plankton); the latter is expected to have higher dispersal potential because larvae can potentially stay in the ocean currents for longer and

thus travel further. However, long-distance dispersal of planktotrophic and lecithotrophic larvae can be highly dependent on larval swimming capacity and behaviour (Nanninga & Manica 2018).

To complete the process of dispersal, larvae must navigate across the seascape matrix (termed the 'transience' step by Cayuela et al. 2018) and conclude in successful settlement and recruitment into the new population. Though, post-transience there are several factors that can influence settlement and recruitment. For example, long-distance dispersers run the risk of being carried beyond suitable habitat or beyond tolerable environmental conditions (not to mention an increased risk of predation), particularly if they originated from a location at the periphery of a species' range. Moreover, if the seascape contains a mosaic of suitable habitats then the likelihood of larvae finding an adequate habitat patch decreases compared to seascapes with continuous habitat patches; such information is extremely valuable for conservation managers when protecting species that are found in both types of habitats. Larvae also have to deal with biotic pressures immediately after settlement, such as predation and competition within and between species. For instance, in areas where there is high competition for space and resources, only the fittest individuals and species will survive and proliferate. These factors constantly interact and contribute to shaping the ecological structure of the community, but also influence intraspecific recruitment, evolutionary processes, and connectivity within and between spatially discrete populations (Cowen & Sponaugle 2009).

1.1.4 Genetic connectivity

Genetic connectivity is defined by Lowe & Allendorf (2010) as the degree to which gene flow affects evolutionary processes within populations. From an ecological perspective, this definition implies that individuals must disperse to a new population and successfully contribute their genes into the local gene pool to facilitate genetic connectivity. As a result, genetic methods of assessing connectivity represent a form of realised (or effective) connectivity, whereby only the contributions of dispersers that survive and successfully reproduce are usually quantified. This means, however, that genetic methods do not always account for immigrants that join a population but do not reproduce.

Demographic connectivity refers to how the number of exchanges between populations (via immigration and emigration) affects population growth and vital rates (Lowe & Allendorf 2010). Alternatively, it has been defined as the movement of individuals between populations, the extent of which is large enough to be

demographically significant, where 'significance' is context dependent (Leis 2006). Unlike genetic connectivity, demographic connectivity is a function of the relative contribution of net immigration (immigrants – emigrants) to total recruitment (local recruitment + net immigration). In expanding populations, therefore, net immigration can be very high but still only represent a small proportion of total recruitment and, vice versa, when populations are declining, net immigration can be low but represent a large proportion of total recruitment (Lowe & Allendorf 2010). The latter has been presented as the 'rescue effect' for populations that are near extinction because critically they rely on immigrant subsidy for persistence (Brown & Kodric-Brown 1977). This demonstrates the importance of demographic connectivity for maintaining stability in populations with low local recruitment (Runge et al. 2006) and for (re)colonising unoccupied habitat patches in meta-population systems (Hanski 1998).

The level of exchange required to maintain demographic stability is in orders of magnitude higher than to maintain genetic homogeneity (Lowe & Allendorf 2010). As few as ten effective migrants per generation may be enough to maintain drift connectivity, that is, sufficient gene flow to preserve similar allele frequencies between populations (Lowe & Allendorf 2010). From a conservation perspective, a reduction in demographic connectivity (i.e. decrease in immigrant subsidy) can be cause for concern because it may reduce population viability; similarly, from a fisheries management perspective, insufficient demographic connectivity is equally concerning because it may result in decreases in yield. This may impact the population (or stock) persistence, which is of central importance to managers of marine reserves and resources, and which may not be fully answerable with genetic data, unless combined with other data such as population growth rates, movement behaviour, reproductive success or biophysical modelling (Lowe & Allendorf 2010; Breusing et al. 2016). However, genetic connectivity allows insights into the degree of realised connectivity between two populations, which can also be very useful for managers who are interested in the contribution of immigrants that survive, reproduce and add to the local gene pool. Furthermore, gene flow can play an important role in reducing inbreeding and purging deleterious mutations, as well as spreading advantageous alleles and introducing new adaptive variants into the population (Frankham 2015). In fact, as few as one migrant per generation may be enough to reduce the harmful effects of inbreeding, termed inbreeding connectivity by Lowe & Allendorf (2010), thereby maintaining fitness by counteracting the loss of genetic diversity (Frankham 2015).

Estimating genetic connectivity

Gene flow between spatially discrete populations can be estimated using indirect and direct approaches (Gagnaire et al. 2015), both of which rely on the development of adequate genetic markers (section 1.2.2). Indirect approaches focus on estimating the amount of genetic divergence between populations and, in contrast, direct approaches attempt to detect migrants by using multi-locus genotypic information to assign individuals to their location or parents of origin (Lowe & Allendorf 2010; Gagnaire et al. 2015). These direct approaches are, by design, quite similar to some non-genetic methods (section 1.2.1) that are used for identifying immigrants among populations.

Indirect approaches typically use genetic indices, such as F_{st} (Weir & Cockerham 1984) and D (Jost 2008), and population structure analyses to estimate the degree of genetic divergence among populations. These methods are sometimes based on a number of assumptions (e.g. Hardy-Weinberg equilibrium) relating to population dynamics (e.g. constant population size), life history (e.g. non-overlapping generations) and the contribution of evolutionary forces (e.g. negligible effect of selection) (Gagnaire et al. 2015), which are infrequently met in natural systems. Moreover, it is important to be aware that low levels of genetic differentiation (suggesting migration is above the threshold required for genetic connectivity) do not guarantee demographic connectivity, a phenomenon recently defined as 'crinkled connectivity' (Ovenden 2013). To further complicate interpretations, populations can be genetically similar in the absence of gene flow due to large effective population sizes (N_e), which can mitigate the influence of genetic drift; unfortunately, however, accurately estimating N_e in marine organisms has been notoriously difficult (Hare et al. 2011). Nevertheless, these indirect approaches have proved vital for studying the processes of gene flow, drift and selection in natural populations (Hellberg 2009).

Direct approaches can be broadly divided into methods of population assignment and parentage analysis (Manel et al. 2005); both approaches make fewer assumptions than indirect approaches but require sound knowledge of the species distribution (Gagnaire et al. 2015). Individual assignment assigns an individual to a population or location in which their genotype has the highest probability of occurring. Although devoid of model-based assumptions, this approach is highly sensitive to the degree of genetic differentiation between populations as the accuracy of assignment is proportional to the genetic differentiation between putative populations within a species (i.e. decreased genetic differentiation equals decreased accuracy) (Christie et al. 2017). This has been a problem in many studies of benthic marine invertebrates because they

often exhibit globally weak genetic differentiation (Benestan et al. 2015); however, this limitation *may* be alleviated with the advent of genomics (section 1.2.3), which promises more powerful markers and higher resolution for detecting genetic differences between sampling sites (Allendorf et al. 2010). In contrast to individual assignment, parentage analyses assign an individual to their biological parents based on their observed genotypes (Jones et al. 2010). Although relatively insensitive to the amount of genetic differentiation, parentage assignment is highly sensitive to the proportion of potential parents sampled (Christie et al. 2017). Therefore, for accurate assignment, a significant proportion of potential parents must be sampled, which is frequently logistically very difficult to achieve in the marine realm (but see Saenz-Agudelo et al. 2009; Harrison et al. 2012; Williamson et al. 2016).

Evolutionary processes influencing population structure

Investigating genetic diversity and population genetic structure are often the first analyses to be undertaken in a conservation or fisheries genetics study. As mentioned previously, this enables an indirect assessment of genetic connectivity by providing insights into patterns of gene flow, but these analyses also enable the study of other evolutionary processes acting on populations, such as drift and selection (Hellberg & Burton 2002). For example, when there are restrictions to gene flow, population allele frequencies can diverge over time due to drift (the random sampling of alleles from generation to generation); therefore, with the assumption of random mating, an allele could become fixed or completely purged from a population by chance. Moreover, drift is stronger in small or bottlenecked populations because the sampling variance is greater when the N_e is smaller (Charlesworth 2009). Populations can also diverge when strong natural selection favours a particular mutation that increases the fitness or survivorship of the carriers, resulting in the allele sweeping to fixation in that particular population (Nielsen 2005). The absence of gene flow has been assumed to be favourable for local adaptation because gene flow can swamp locally adapted alleles (Morjan & Rieseberg 2004). However, contrary to this, it is now widely acknowledged that local adaptation can develop in high gene flow scenarios (e.g. Cure et al. 2017; Diopere et al. 2017), particularly when individuals selectively disperse towards habitats that inherently maximise their fitness (Jacob et al. 2017). For studies where the primary goals are to assess inbreeding, N_e or connectivity, neutral markers are predominantly used because these markers are assumed to be driven by the interacting processes of gene flow and drift, and not selection (in which adaptive markers are more informative). However, it is important to note that a

recent barrier to gene flow may not be detectable in some markers because there has been insufficient time for the allele frequencies to diverge by drift (Hedgecock et al. 2007). This could be circumvented to some degree by also using a direct approach of inferring connectivity (i.e. population or parentage assignment) or by collecting the relevant data needed to explore demographic connectivity.

1.2 Measuring connectivity

Measuring connectivity is not trivial in the marine environment as the distribution and migratory pathways of marine organisms are concealed beneath the surface of the oceans, hidden from the human eye. Nevertheless, there are a number of different techniques used to measure and quantify connectivity within and between populations in marine systems. The choice of which technique(s) to employ depends on the study organism and the objectives of the research project, but, typically, both logistics and finance must be considered as these factors can be major limiting factors in the scope of marine connectivity studies. As alluded to in the previous section, there are two types of approaches to measuring connectivity: direct and indirect methods. Direct methods rely on recording movement by visual observations or by tracking the organism through tagging (genetic, chemical or physical) or assignment techniques. In comparison, indirect methods infer movement and connectivity but are not constrained by the need to recover tags or recapture the animal, and knowledge of the natal origin is not required. The remainder of this section considers the approaches used –both non-genetic and genetic– to infer connectivity, with the majority of the section focusing on the genetic markers used as these are primary techniques used in this thesis.

1.2.1 Non-genetic methods

Direct observations are the most simplistic method of studying the movements of an organism. This typically involves a capture-mark-recapture approach, whereby the organism is caught, uniquely marked and released, and when the organism is recaptured its patterns of movement can be assessed. This method can be effective for monitoring the population sizes and movements of larger animals; however, a major drawback is that the organism must be recaptured. Furthermore, this approach can only tell us where an individual was at a certain time at a certain place – it does not tell us how the individual got to that place.

To determine the exact route of travel, novel techniques using electronic tagging devices have been deployed which enable real-time tracking of an organism, both horizontally (i.e. across seascape) and vertically (i.e. depth) (Cooke et al. 2013).

For example, satellite telemetry has been used to explore the three-dimensional movements of leatherback turtles (*Dermochelys coriacea*) in the Atlantic and has provided insights into the spatial movements and diving behaviour of this species (Hays et al. 2004). However, although electronic tagging has been extensively used in tracking marine vertebrates (Costa et al. 2012; Hazen et al. 2012), it is still in its infancy for marine invertebrates (Fossette et al. 2016). Moreover, despite the technological advances in reducing the size of tags and increasing the amount of data gained per device (Cooke et al. 2013), it is still impossible to track larvae composing the meroplankton using this technique. Pelagic larvae of benthic marine organisms typically range from 20 microns (microplankton) to two centimetres (macroplankton) in size, while the smallest satellite tag to-date known to provide useful data is $\sim$1.6 grams (Fig. 3), which was used to track critically endangered spoon-billed sandpipers (www.saving-spoon-billed-sandpiper.com). Therefore, for marine organisms that have a pelagic larval dispersal phase, other methods are needed to measure connectivity.

Figure 3: One of the smallest satellite tags developed to-date, used to track critically endangered spoon-billed sandpipers (*Eurynorhynchus pygmeus*) (Photo by Chris Kelly).

Tracking individuals using geochemical differences across water bodies is another empirical non-genetic method used to study movement and connectivity in marine and freshwater organisms (Thorrold et al. 2007). These techniques rely on using naturally occurring geochemical tags (e.g. otoliths in fish and stratoliths in some invertebrates), in which movement across different environmental gradients

(e.g. temperature) leaves a distinct detectable signature in the calcified structures of these geochemical tags (Thorrold et al. 2007). This method has been adopted in several fields of study to examine, for example, temporal and spatial patterns of connectivity and settlement in mytilid mussels (Fodrie et al. 2011; Gomes et al. 2016), and population connectivity of coastal fishes (Fodrie & Herzka 2013; Williams et al. 2018). However, little or no information about connectivity can be extracted when recruits originate from locations with similar environmental signatures (i.e. similar sea chemistry) (Fontes et al. 2009). Moreover, this method is only applicable to organisms that possess geochemical tags.

1.2.2 Genetic markers: introduction

Genetic markers (or molecular markers) are any gene, DNA sequence or other molecular-based unit that can be used to study genetic variation at the individual, population or species level. They are naturally occurring markers present in every individual and are, therefore, very useful for studying population structure and inferring connectivity. The origin of genetic markers dates back to the 1970s, during which time allozymes were first used to investigate genetic variation and detect siblings in a range of species (Allendorf 2017). By 1980, following the development of polymerase chain reaction (PCR) amplification, mitochondrial DNA (mtDNA) genes were being sequenced to explore phylogeography and population structure (Allendorf 2017). The next few decades saw technological advances give rise to more variable and increased numbers of loci (e.g. microsatellites and SNPs), followed by the genomic revolution in the latter 2000s (Metzker 2010; Allendorf et al. 2010; Davey et al. 2011). A comprehensive discussion of genetic markers and their uses in molecular ecology has been presented in several renowned review papers (Sunnucks 2000; Schlötterer 2004; Hellberg & Burton 2002; Allendorf 2017); however, some of the markers most commonly used for studying historical and contemporary population biology, and pertinent to the research presented in this thesis, are discussed below.

mtDNA

Mitochondria have their own genome separate from nuclear DNA. The robustness of mitochondrial DNA (mtDNA) as template DNA for subsequent PCR amplication, the variability of particular genes composing mtDNA and the general lack of recombination have contributed to the popularity of mtDNA over the last few decades (Rowe et al. 2017). Moreover, the conserved arrangement of genes in mtDNA across the animal kingdom has led to the development of universal

primers, which are now readily available. In most animals mtDNA is maternally inherited because sperm mitochondria are usually destroyed post-fertilisation (Rowe et al. 2017). This trait of uniparental inheritance means mtDNA can be treated as a single haplotype (with all sites sharing a common genealogy), allowing individual lineages to be tracked over time, which has useful applications in phylogeography (Avise et al. 1987; Hickerson et al. 2010; Puebla 2018). However, this effectively means that these are single-locus markers, and because only the maternal lineage is investigated, there can be erroneous conclusions about genetic breaks in species where there is male-biased dispersal. For example, in scenarios where females do not disperse or there is philopatry in females but not in males (e.g. loggerhead sea turtles, Casale et al. 2002), mtDNA haplotypes may be highly differentiated when in fact there is male-mediated gene flow (Stiebens et al. 2013). Another caveat of mtDNA is that copies of mtDNA can be translocated onto the nuclear genome, known as mtDNA pseudogenes or nuclear mitochondrial DNA segments (numts) (Hazkani-Covo et al. 2010). These pseudogenes are non-functional and continue to evolve independently of mtDNA, which can introduce problems when primers unintentionally amplify numts in addition to the desired mtDNA fragment (Calvignac et al. 2011). mtDNA has long been considered a neutral marker, but these assumptions are starting to be questioned, with one review suggesting mtDNA is far from neutrally evolving (Galtier et al. 2009). Nevertheless, mtDNA markers still have useful applications as a first look into genetic structure because drift is stronger compared to nuclear markers due to the lower N_e (Puebla 2018), but the addition of complementary nuclear markers is generally thought to overcome the caveats of only using mtDNA markers for interpreting genetic patterns.

Microsatellites

Microsatellites are short tandem repeats of typically 2-6 bp (base pairs). They are codominant markers, meaning homozygotes and heterozygotes can be distinguished; this allows locus and population allele frequencies to be calculated. Microsatellites are generally considered neutral markers because the repeat variations are thought to be non-functional; though, few have been thoroughly assessed and there are some known exceptions, such as the existence of some loci associated with human neurodegenerative diseases (Vieira et al. 2016). Due to their short repeat motif, microsatellite loci have high mutation rates, which can result in a large number of alleles of different sizes, making these markers highly informative for exploring genetic variation across individuals and populations (Schlötterer 2004). However, microsatellites have complex mutation behaviour and

are generally species-specific (Schlötterer 2004); moreover, difficulties in cross-calibrating microsatellite allele sizes between sequencing platforms and laboratories (e.g. Ellis et al. 2011) has limited their use in broad-scale studies. Nevertheless, they have been, and still are, widely used in marine and terrestrial population genetics, although their usage has waned as new technology permits the high-throughput genotyping of other polymorphic markers.

Single nucleotide polymorphisms

Single nucleotide polymorphisms (SNPs) are single base changes at a position in the DNA sequence (Fig. 4). SNPs have become more popular over the last decade because they are abundant across the genome, have a simple mutation model and are eligible for high-throughput screening and automation (Seeb et al. 2011; Helyar et al. 2011). It is also relatively straightforward to calibrate SNPs among laboratories and to assemble both spatial and temporal datasets from multiple studies (Helyar et al. 2011). Although a single biallelic SNP locus contains less information (i.e. less allelic diversity per locus) than a microsatellite locus, individual SNPs can segregate strongly among populations (Helyar et al. 2011). Some reports suggest that 100 neutral SNPs have approximately the same discriminatory power to detect population structure as 10-20 microsatellites (Kalinowski 2002). However, the most informative SNPs (i.e. those that show the greatest allele frequency variation among populations) have the potential to rival or even exceed the power of microsatellite markers (Helyar et al. 2011). In combination with the rapid rise of genomics, this has led to an increase in the development of small panels of informative SNPs for a variety of marine and terrestrial species (e.g. salmonids, Meek et al. 2016; crustaceans, Jenkins et al. 2018b; and molluscs, Jiao et al. 2014).

1.2.3 Genetic markers: the genomic revolution

The advent of next-generation sequencing (NGS), also called massive parallel sequencing, has paved the way for large amounts of sequence data to be generated at more affordable costs (Ellegren 2008; Metzker 2010; Mardis 2011). This introduced new bioinformatics challenges (Pop & Salzberg 2008); however, year upon year increases in the read lengths generated by sequencing platforms (Quail et al. 2012; Pillai et al. 2017) and improvements in algorithms and computational power, have somewhat alleviated these challenges. With the enormous amounts of data that can now be gleaned from the genomes of model and non-model organisms, it has been said that we are in a genomic revolution

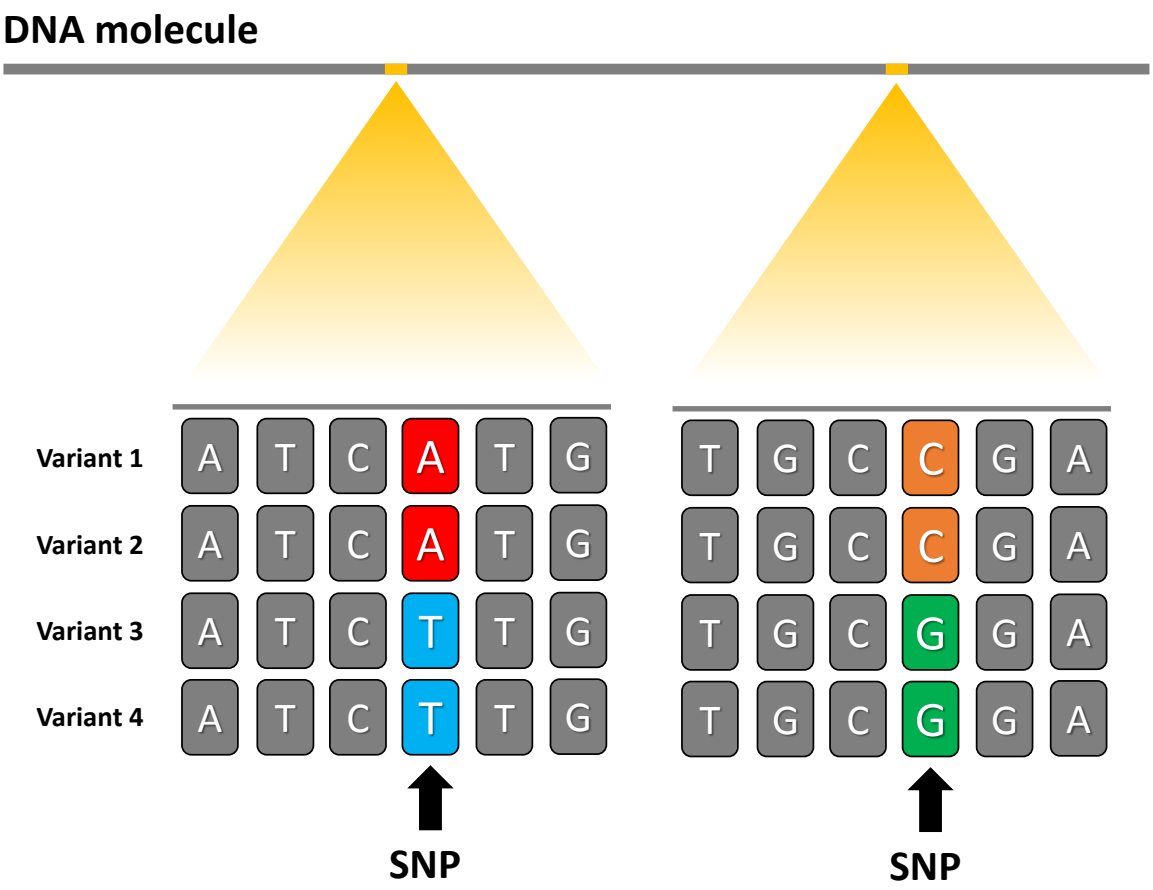


Figure 4: Single nucleotide polymorphisms (SNPs) at two positions in a DNA molecule. A SNP is a single base change at a position in a DNA sequence.

(Ellegren 2014), and with the continual advancement of sequencing technology, such as third-generation single-molecule real-time sequencing (e.g. Pacific Biosciences, Oxford Nanopore Technologies), it seems that the revolution is not slowing down just yet.

NGS technology has enabled the development of new methods of marker isolation for studies of population genetics, adaptation and conservation biology (Andrews & Luikart 2014). For example, exon-capture (Hodges et al. 2007), one such method made possible by NGS, targets and isolates SNPs from the exome, thereby permitting the study of protein-coding genomic regions (reviewed in Mamanova et al. 2010). Another method, next-generation cDNA sequencing (RNAseq), makes it possible to isolate SNPs (Barbazuk & Schnable 2011) or to sequence entire transcriptomes from almost any individual or tissue (Ozsolak & Milos 2011; Todd et al. 2016). However, one of the most popular methods that emerged from NGS is restriction-site associated DNA sequencing (RADseq) (Miller et al. 2007; Baird et al. 2008), in which short regions of DNA adjacent to restriction enzyme cut sites are sequenced. This is a type of reduced-representation

sequencing (RRS) approach, in which a subset of markers are isolated from across the genome, the number of which will depend upon the size of the genome and the cutting frequency of the restriction enzyme used (Davey et al. 2011).

Since the conception of RADseq, a vast number of similar methods have been introduced, some of which have only minor tweaks or modifications to the traditional RADseq approach, while others have more pronounced differences (Campbell et al. 2018). For example, nextRAD was recently developed by some of the original authors of RADseq (Baird et al. 2008; Etter et al. 2011) but uses a combination of selective primers and transposomes to cut DNA instead of restriction enzymes (Fu et al. 2017). One of the main advantages of RADseq and some of its derivatives is that no prior knowledge of the genome is required to perform the techniques (Rowe et al. 2011), which is a massive advantage for working with non-model species (species without a reference genome). However, if a reference genome is available for the study species (or a closely related species), NGS reads can be aligned to the reference genome; this can have very useful benefits for RADseq studies with non-model species such as: (i) improved assembly and identification of SNP loci (by reducing potential effects of sequencing error); (ii) enhanced ability to filter paralogous or repetitive sequences and remove non-target DNA (contamination); (iii) allowing the physical position of loci to be considered (advantagous for mapping studies); and (iv) increased statistical power to detect genomic regions of interest, for example regions under divergent selection between populations (Andrews et al. 2016). Another major advantage of these techniques is that thousands to tens of thousands of genome-wide SNPs can be discovered from across the genome in a single sequencing run (Andrews et al. 2016). In comparison, typical microsatellite development usually consists of tens of markers, and there is no way of knowing if these microsatellites are distributed evenly across all of the chromosomes. In studies of population genetics and local adaptation, the ability to potentially sample across all chromosomes is highly advantageous because it offers more opportunities to find informative loci (neutral or adaptive). Of course, in non-model organisms, there is also no way of knowing whether the SNPs discovered from a RADseq approach are distributed evenly across the chromosomes; however, based on our knowledge of how restriction enzymes operate, and the sheer number of SNPs discovered using these approaches, it is likely that a large proportion of chromosomes will be represented, maximising the chances of finding some interesting loci.

RADseq approaches have now been applied to many different research fields to answer a variety of questions about ecology and evolution. For example, they have been used in studies of species identification (Maroso et al. 2018), species

delimitation (Pante et al. 2015a; Herrera & Shank 2016), hybridisation (Faust et al. 2018), stock management (Mullins et al. 2018), population assignment (Drinan et al. 2018), phylogeography (Emerson et al. 2010), local adaptation (Harrisson et al. 2017), connectivity (Van Wyngaarden et al. 2017; Xuereb et al. 2018) and growth-related traits (Yu et al. 2018). One of the areas in which RADseq approaches have become particularly useful is in accurately resolving population structure at broad- and fine-scales. In marine species, determining population structure has been challenging due to typically weak genetic differentiation and the limited resolution offered by traditional molecular markers (Benestan et al. 2015). However, RADseq approaches enable the discovery of numerous genome-wide markers which maximises the power to detect subtle genetic differences among populations (Funk et al. 2012). Such an approach has allowed researchers to resolve fine-scale population structure in range of marine species including great scallops (Vendrami et al. 2017), emperor penguins (Younger et al. 2017), American lobsters (Benestan et al. 2015), staghorn coral (Drury et al. 2017) and starlet sea anemones (Reitzel et al. 2013). On the other hand, it has also confirmed the existence of no population structure across certain spatial scales (e.g. Everett et al. 2016; Pérez-Portela et al. 2018), which is still an equally important finding for marine management as it implies genetic connectivity across the geographical scale of the samples analysed.

RADseq and genomics have also contributed to the discovery of 'outlier' markers, which are candidate markers potentially under strong drift or divergent selection (Lotterhos & Whitlock 2015). From a connectivity perspective, these outlier markers typically have greater power to differentiate populations, which has promising applications for inferring connectivity using population assignment approaches (Gagnaire et al. 2015). Indeed, the incorporation of gene-associated markers in assignment has already proven to be incredibly useful for fisheries management, where these markers are used as tools to help tackle illegal fishing (Martinsohn & Ogden 2009; Nielsen et al. 2012). From a conservation perspective, outlier markers have the potential to revolutionise the delineation of conservation units (section 1.3.1) by identifying adaptive diversity in protected species (Funk et al. 2012). Genomics appears to have been widely accepted by academic conservationists, with many recent review papers dedicated to discussing the opportunities provided by conservation genomics (e.g. Allendorf et al. 2010; Benestan et al. 2016a; Flanagan et al. 2017; Barbosa et al. 2018); however, a major challenge is the translation and integration of genomic data into conservation practice (Shafer et al. 2015; Garner et al. 2016).

1.2.4 Biophysical modelling

All previous methods described in this section provide an empirical assessment of connectivity, each with their own merits and caveats. As an alternative, powerful biophysical models have been developed to assess marine connectivity; these utilise biological and hydrological data to simulate larval dispersal across seascapes (Metaxas & Saunders 2009). Biophysical models use outputs of ocean models (e.g. current velocities, directions, etc.) as inputs to particle tracking algorithms that track individual particles (larvae) from a starting point to settlement (Cowen 2006).

However, to accurately predict larval dispersal using these approaches, incorporating accurate biological parameters is critical. For example, the length of time spent drifting is usually determined by an organism's PLD, which can be estimated from laboratory or (preferably) field studies (Metaxas & Saunders 2009). However, many other factors can influence larval dispersal that could be integrated into the model, such as mortality (Treml et al. 2015), spawning time and periodicity (Kough & Paris 2015), salinity and temperature (Lárez et al. 2000), and settlement likelihood and larval behaviour (Treml et al. 2012). Simulations that are performed with the incorporation of such species-specific data are usually termed individual-based models (IBMs) (DeAngelis & Grimm 2014).

Biophysical models provide insight into the larval dispersal potential and demographic connectivity of the species under study (Thomas & Bell 2013). They have also been used to estimate dispersal kernels, defined as the spatial probability distribution of dispersal distances based on repeated events, which theoretically capture the temporal variability of larval dispersal (Siegel et al. 2003; Cowen & Sponaugle 2009). More recently, biophysical modelling has been combined with genetic studies to explore seascape genetics or seascape genomics (Riginos et al. 2016; Selkoe et al. 2016). As with other naming conventions (i.e. population genetics/genomics, conservation genetics/genomics, etc.), seascape genomics is generally analogous to seascape genetics, with the exception that seascape genomic studies genotype markers using high-throughput genomic techniques such as RADseq. Seascape genetic studies incorporate the fields of ecology, oceanography and geography to explore and interpret patterns of marine connectivity (Christie et al. 2010; Selkoe et al. 2016). This approach has provided novel insights into spatial patterns of population structure and connectivity in many species including seagrass (e.g. Jahnke et al. 2018), crustaceans (e.g. Thomas & Bell 2013; Benestan et al. 2016b) and molluscs (e.g. Breusing et al. 2016; Sandoval-Castillo et al. 2018), and is likely to have many applications in

fisheries science, conservation and marine reserve design (Selkoe et al. 2016; Mertens et al. 2018).

1.3 Marine conservation

The main aim of marine conservation is to restore and/or protect marine ecosystems to preserve biodiversity and to avoid overexploiting marine resources. One of the first world-wide political movements to recognise the importance of conserving natural environments was founded in 1992: the Convention on Biological Diversity (CBD); this was was initially signed by 150 countries (currently 182 countries and the European Union) (Ten-Kate 2002). The objectives of the CBD are to (i) conserve biological diversity, (ii) ensure the sustainable use of its components, and (iii) enable fair and equitable sharing of benefits arising from the use of genetic resources (Ten-Kate 2002).

The CBD movement coincided with the establishment of the Oslo/Paris (OSPAR) Convention, which was inaugurated to protect the marine environment of the northeast Atlantic. The OSPAR Convention is a collaboration of fifteen governments and the European Union (EU) (also called contracting parties) whose purpose is to develop policy and international agreements to safeguard the OSPAR Maritime Area in the northeast Atlantic (Fig. 5). In 1994, the OSPAR Maritime Area was divided into five regions for assessment and monitoring purposes: Arctic Waters (region I), Greater North Sea (region II), Celtic Seas (region III), Bay of Biscay and Iberian Coast (region IV), and Wider Atlantic (region V) (OSPAR Commission 2013). Legislation within and across countries has followed the OSPAR Convention which empowers governments to deliver towards the OSPAR agreement. For example, the European Union introduced the EU Marine Strategy Framework Directive (MSFD) in 2008 and the UK introduced the Marine and Coastal Access Act in 2009. The MSFD requires that each member state establishes 'coherent and representative networks' of Marine Protected Areas (MPAs) by 2020. Indeed, from a UK government perspective, its recent movement towards protecting marine environments in UK waters also contributes to its commitments to the CBD and other international agreements, thereby satisfying both politicians and conservationists.

1.3.1 Conservation units

Delimiting conservation units (CUs) is an essential first step for managers and policymakers so that they know the boundaries of the populations they are trying to conserve (Funk et al. 2012). Defining such boundaries allows the status of a

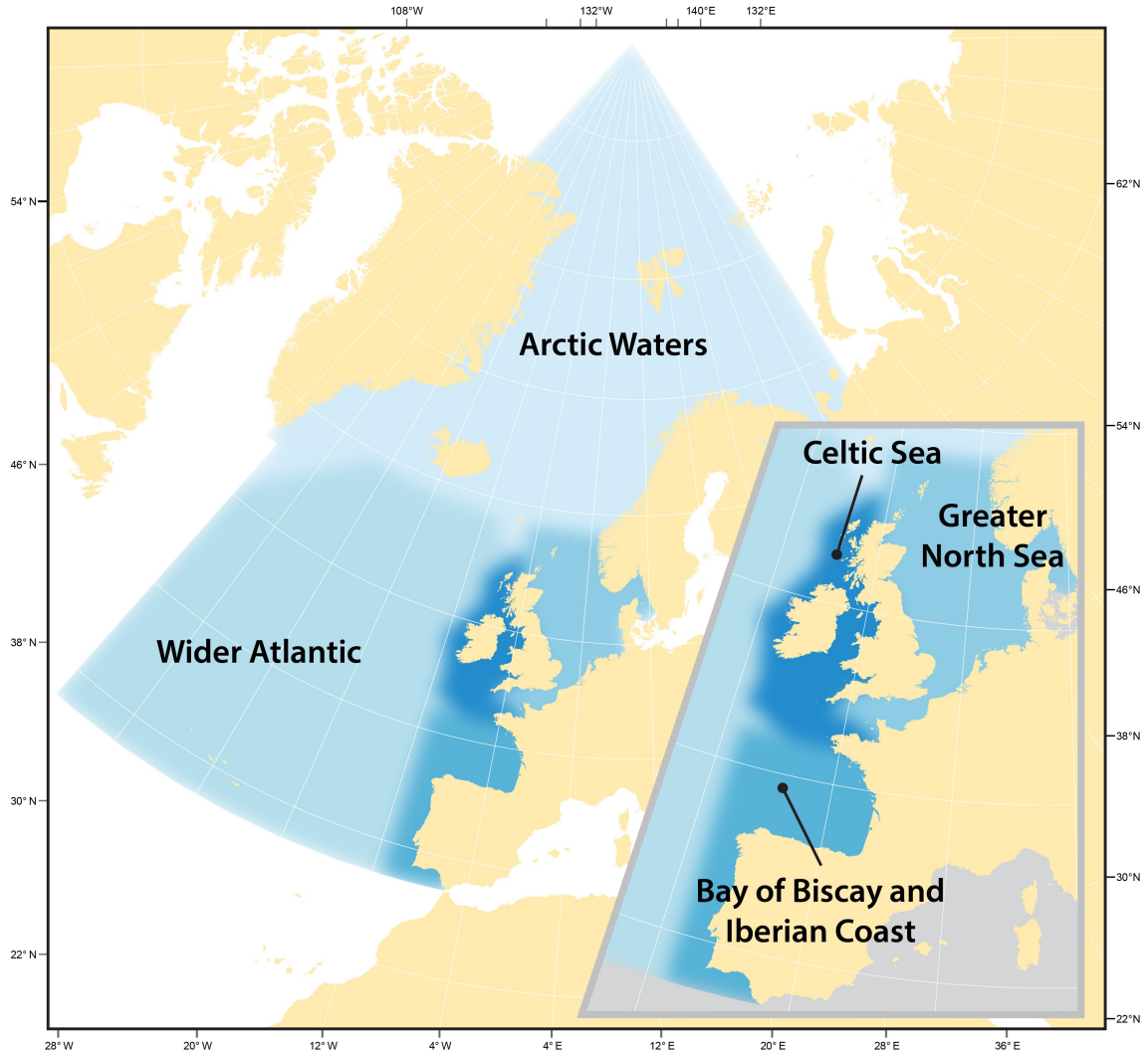


Figure 5: The OSPAR Maritime Area. This area is divided into five regions for assessment purposes: Arctic Waters (region I), Greater North Sea (region II), Celtic Seas (region III), Bay of Biscay and Iberian Coast (region IV), and Wider Atlantic (region V).

population to be assessed, which can inform the development of a management strategy. However, defining populations can be extremely challenging in habitats that are continuous over larges spaces (Waples & Gaggiotti 2006), which can be a prevalent feature of marine environments. Nevertheless, CUs are still relevant for marine species, particularly for the management of fishery stocks and for the protection of vulnerable species (Heyden et al. 2014).

Conservation units have generally been discussed in terms of evolutionary significant units (ESUs), management units (MUs) and, recently, adaptive units (AUs) (Funk et al. 2012; Barbosa et al. 2018). Definitions of ESUs are plentiful (Box 1, Funk et al. 2012), though overall there is general agreement that an ESU represents a group of conspecifics that exhibit high genetic and ecological distinctiveness (Ryder 1986; Funk et al. 2012), often due to allopatric or adaptive

divergence (Moritz 1994). Conserving ESUs is, therefore, a high priority for management because the maintenance of different ESUs will maximise the evolutionary potential of a species to adapt to environmental change (Funk et al. 2012). Populations that are demographically independent (i.e. where population growth rate is dependent on local recruitment and not immigration) are typically classed as MUs (Moritz 1994). Identifying these units can be particularly important in fisheries for delineating stocks with distinct population growth rates and demography (Palsbøll et al. 2007; Heyden et al. 2014). Whereas ESUs typically consider all genetic variation and MUs only consider neutral genetic variation (Funk et al. 2012), AUs specifically describe the adaptive differences between populations, which can be very important for prioritising conservation resources and for deciding which individuals to use as sources for supplementing depauperate populations (Moritz 1999). For example, supplementing a focal population with genetically incompatible sources (e.g. individuals adapted to a very different environment) may lead to outbreeding depression (Frankham et al. 2011). For restocking programs, this knowledge is crucial so that hatchery managers ensure that the juveniles bred from their facility are compatible with the target population or area being stocked (Ward 2006).

Allocating and prioritising conservation resources is not trivial. For example, if resources were only available to conserve one or two populations, managers must weigh-up which populations are of highest value to the conservation of a species overall. This is where CUs and other relevant data sources, such as measures of genetic diversity, are central for conservation managers. For instance, stepping stone sites (and associated populations) that link isolated populations can be vital for long-distance dispersal, connectivity and range expansions (Saura et al. 2014); therefore, these areas may be prioritised to maintain linkages between ecosystems. In the marine environment, such a prioritisation approach is often used to help design networks of Marine Protected Areas.

1.3.2 Marine Protected Areas

Marine Protected Areas (MPAs) are areas of sea or ocean designated to protect habitats or species, or both. At the time of writing, over 11,000 MPAs have been designated globally to protect the world's oceans (www.mpatlas.org). This equates to around 3.7 % of all oceans on Earth being protected. However, not all of these MPAs are currently implemented, well-managed or enforced, which is often because of capacity shortfalls associated with inadequate staff and finances (Gill et al. 2017).

MPAs vary greatly in size and in the stringency of protection, ranging from large

no-take zones (NTZs) where typically no activity is permitted, to smaller localised designations that may prohibit a specific recreational activity or the fishing of a particular species. Research has suggested that the conservation benefits from MPAs increase exponentially when MPAs: (i) are NTZs, (ii) are large ($>$100 km^2), (iii) are well regulated and enforced, (iv) have been establish for a long time ($>$100 years), and (v) are isolated by deep water or sand (Edgar 2004). However, although large MPAs make substantial contributions towards the CBD's Aichi Target 11 (protecting 10 % of coastal and marine areas by 2020), their contribution to marine conservation has been debated (Davies et al. 2017; O'Leary et al. 2018). The negativity has mainly stemmed from a lack of understanding of their actual benefits for conserving biodiversity, and that these large-scale MPAs are (i) driven predominantly by political targets, and (ii) situated typically in remote areas where there are minimal threats and where establishment is easier (Leenhardt et al. 2013; Devillers et al. 2015). On the other hand, many advocate that they are essential for protecting wide-ranging or circumtropical species such as seabirds and tuna (Young et al. 2015), and that they are important for maintaining pristine areas (e.g. the Chagos Archipelago, Sheppard et al. 2012) or for capturing habitat shifts associated with climate change (Toonen et al. 2013). Moreover, it seems logical to create legislation to protect these remote areas now before they are potentially targeted for exploitation or habitation in the future (Toonen et al. 2013). Irrespective of these contrasting viewpoints, it is inevitable that large-scale MPAs will indeed become a prominent feature in our oceans in the coming years as governments seek to meet global conservation targets.

Despite the recent furore surrounding large MPAs, small MPAs still remain critical for marine conservation (Toonen et al. 2013), particularly in coastal areas within the jurisdiction of individual governments. For example, in the northeast Atlantic, individual governments have designated a number of MPAs within their exclusive economic zone (EEZ) with the aim of establishing a MPA network to protect local features (i.e. habitats and species) and, of course, to satisfy their commitments to international agreements. A network of MPAs is thought to collectively deliver more benefits to biodiversity than individual, unrelated MPAs (Foster et al. 2017). Countries have full rights to manage the marine resources in their EEZ, which extends 200 nautical miles (370 km) from the coastline, allowing each government to designate MPAs anywhere within their territorial waters. However, unless these MPAs have well-thought-out management and monitoring plans, with adequate funding and staff to carry out the plans (Álvarez-Fernández et al. 2016), they are essentially just polygons on a map/chart. The challenge for local and regional marine managers is, therefore, how to strategically allocate

available resources in a way that maximises protection of marine biodiversity, while also considering the impact to stakeholders that are involved, interested and/or affected by the MPA (Dehens & Fanning 2018). Simultaneously, at the government level, a further challenge is how to demonstrate that MPAs are not simply designated randomly but are distributed purposefully across territorial waters such that they are an 'ecologically coherent' network (Ardron 2008a), a requirement for contracting parties of the OSPAR Convention and EU member states.

1.3.3 Ecological coherency

As part of the OSPAR Convention, contracting parties agreed to establish an ecologically coherent well-managed network of MPAs, initially by 2010 (Ardron 2008a). Yet, at the time the meaning of ecological coherency or how to assess whether a network is ecologically coherent or not was not explicitly stated (Ardron 2008b). Since then, OSPAR reports have generally defined an ecologically coherent MPA network as a network that considers Adequacy, Viability, Representation, Replication and Connectivity (OSPAR Commission 2007; 2013). Together, these criteria influence and take into account the size of MPAs, the coverage of species and habitats, the spatial distribution of MPAs across biogeographical regions, the number of replicate sites for specific features of interest, and the links between sites at different spatial and temporal scales (OSPAR Commission 2013). However, in 2013, the methods developed for assessing these criteria were still being refined and re-evaluated as the availability of data and knowledge of marine ecosystems increases (OSPAR Commission 2013).

The five principles of the recent 2016 status report of the OSPAR network of MPAs were phrased slightly differently to previous reports: Features, Representativity, Connectivity, Resilience and Management (Fig. 6), but which overall make it easier to interpret the OSPAR principles of ecological coherence (OSPAR Commission 2017). Essentially, Adequacy and Viability have been merged into Features, Resilience encompasses elements of Replication, and Management has been officially added as a focal criterion of ecological coherence (which is a welcome addition due to the emphasis of a 'well-managed' network outlined in the mission statement of the OSPAR Convention). To assess these five principles in 2016, the 'Madrid Criteria' was applied, which was designed to reflect the key network principles, while acknowledging the data limitations associated with target species and habitats, and OSPAR MPA performance (OSPAR Commission 2017). For example, for connectivity, the Madrid Criteria stated that

MPAs must be geographically well-distributed, with a maximum distance between MPAs of 250 km in coastal waters, 500 km in offshore waters, and 1000 km in areas outside national jurisdiction (OSPAR Commission 2017). Other criteria for assessing Representativity, Features and Resilience were also outlined and in 2016 all principles were formally assessed using these criteria. However, the report concluded that although significant progress has been made in developing the network, the OSPAR Maritime Area cannot yet be considered ecologically coherent, with OSPAR citing (once more) that further development of methods to assess ecological coherence is required going forward (OSPAR Commission 2017).

Features
MPAs should be designated in areas that best represent the range of habitats, species and ecological processes in the OSPAR Maritime Area. Proportions of features that should be protected by the MPA network may be higher for particularly threatened and/or declining features.
Representativity
MPAs should protect examples of the same features across their known biogeographical extent to reflect known sub-types. EUNIS Level 3 habitats are stated as a potentially useful way of characterising the OSPAR Maritime Area for the purposes of including biogeographic variation in the network.
Connectivity
In the absence of dispersal data, connectivity may be approximated by ensuring the MPA network is well distributed geographically. Where scientific understanding is further developed, the MPA network should reflect locations where a specific path between identified places is known (e.g. critical areas of a life cycle for a given species).
Resilience
Replication of features in separate MPAs in each biogeographic area is desirable where possible. The appropriate size of a site should be determined by the purpose of the site and be sufficiently large enough to maintain the integrity of the feature(s) for which it is selected.
Management
OSPAR MPAs should be managed to ensure the protection of the features for which they were selected and to support the functioning of an ecologically coherent network.

Figure 6: OSPAR Convention five key principles for assessing ecological coherence of MPA networks. Definitions are taken directly from the OSPAR Commission 2017 report.

1.3.4 Britain's MPA network

The UK is obliged to establish its own ecologically coherent network of well-managed MPAs because of commitments to the OSPAR Convention, the EU MFSD and the CBD. This network will contribute to the OSPAR MPA network but will also satisfy national commitments announced under the Marine and Coastal Access Act 2009 (England and Wales), the Marine Act 2010 (Scotland) and the

Marine Act 2013 (Northern Ireland). As of March 2018, approximately 24 % of marine/estuarine environments around Britain are within MPAs (www.jncc.defra.gov.uk/page-4549). At the time of writing, the network comprised 299 MPAs consisting of: Special Areas of Conservation (SACs; 105) and Special Protected Areas (SPAs; 107) with marine components, Nature Conservation Marine Protected Areas (NCMPAs; 30 in Scotland), Ramsar sites (Isle of Man), and Marine Conservation Zones (MCZs; 56 in England, Wales and Northern Ireland).

Recently, the UK government has focused on augmenting the network with MCZs in England and Wales (Fig. 7). The MCZ project began in 2008 and was co-led by the Joint Nature Conservation Committee (JNCC) and Natural England (NE) with the aim of identifying and recommending candidate sites that fill in gaps in the MPA network, potentially addressing any deficits in Features, Representativity, Connectivity and Resilience. Four regional projects covering southwest England (Finding Sanctuary), southeast England (Balanced Seas), the Irish Sea (Irish Sea Conservation Zones) and the North Sea (Net Gain) were commissioned and, with the support of an independent scientific advisory panel, they submitted their recommendations to JNCC and NE in September 2011. These recommendations were reviewed by JNCC and NE, who submitted 127 candidate MCZs to the Department for Environment, Food Rural Affairs (DEFRA) in July 2012. Subsequently, DEFRA designated 27 MCZs (tranche one, November 2013) and 23 MCZs (tranche two, January 2016) in English waters and one MCZ in Welsh waters (Skomer Island, 2014). A third tranche of 41 MCZs and the addition of new features to 12 existing MCZs is currently under review at DEFRA following a public consultation in June 2018.

The MCZ project has, in the view of some commentators, coincided with a shift from a bottom-up to a top-down approach, with stakeholder engagement now limited to bilateral consultations (Lieberknecht & Jones 2016). The MCZ project has also steered away from its initial focus on broad-scale networks and has instead concentrated efforts on single-feature conservation (Lieberknecht & Jones 2016), such as protecting vulnerable species (e.g. pink sea fans) and key habitats (e.g. intertidal boulder communities). This deviation from a holistic approach has meant that ascertaining whether the UK network satisfies the connectivity principle of the OSPAR Convention is not straightforward. In England and Wales, assessing connectivity of the network has primarily focused on linking discrete habitats (e.g. littoral rock and hard substrata, sublittoral sediment, etc.), such that each habitat is represented by a MPA every 80 km or less (Carr et al. 2014; 2016), the spacing recommended by Roberts et al. (2010) to maintain ecological connectivity. Connectivity for a discrete habitat is deemed sufficient when 40 km buffers drawn

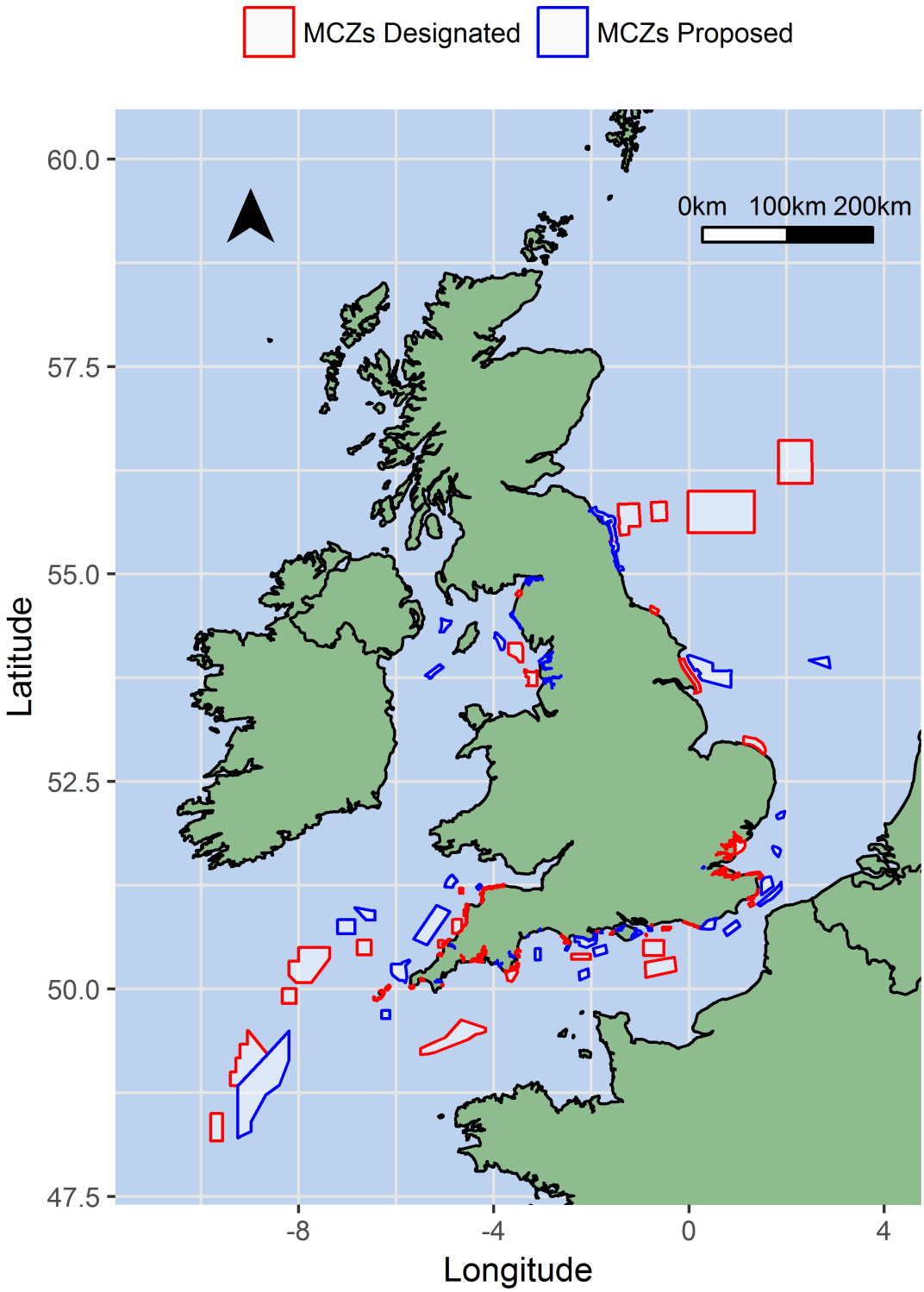


Figure 7: Marine Conservation Zones (MCZs) designated (red outline) and proposed (blue outline) in English waters.

around two adjacent MPAs converge (Carr et al. 2014; 2016). For many benthic marine species, defining a network in this way may be sufficient to maintain connectivity between nearby populations. However, it is important to note that it may not suit all species because, as discussed in section 1.2, connectivity can be influenced by a number of biological and hydrological factors. Moreover, deciding which species to include in assessments of connectivity is not trivial, and compromises are likely to be made in situations for which there are few available dispersal data or there are too few resources to generate novel data.

1.4 Research aims and hypotheses

The overall aim of this thesis was to investigate which taxa may be best suited for assessing genetic connectivity between MPAs, and to assess spatial genetic diversity and connectivity in the species chosen using population genetics and genomics. To do this, a literature survey of population genetic and phylogeographic studies was firstly conducted to explore what has already been documented about the spatial genetic patterns of marine taxa across the northeast Atlantic. In addition to studying genetic patterns across a broad range of taxa, this meta-analysis provided further insights into the biological and methodological information needed to ascertain which taxa may be considered as good candidates for assessing genetic connectivity between MPAs across the British Isles and the wider northeast Atlantic.

Secondly, the knowledge gained from the literature survey facilitated the design of a set of criteria that identified ideal traits of a candidate species for assessing genetic connectivity; subsequently, based on these criteria, this led to the selection of two species for further study. Finally, microsatellite and SNP markers were employed to explore the population genetic structure of these species and these data were used to infer connectivity between MPAs designated within their respective ranges. Specific research questions and hypotheses for each component are outlined below:

Comparative phylogeography meta-analysis

1. Are there common phylogeographic patterns across marine taxa in the northeast Atlantic?

   H_0: Marine taxa have contrasting patterns of phylogeography.

   H_1: Marine taxa have common patterns of phylogeography.

2. Were historic population expansions linked to the Last Glacial Maximum?

H_0: Historic population expansions were not linked to the LGM.

H_1: Historic population expansions were linked to the LGM.

Selecting taxa to assess genetic connectivity between MPAs

Can a set of criteria be designed to assist researchers and managers select appropriate taxa to use as surrogates for assessing connectivity between MPAs?

No hypotheses were tested for this component.

Population genetic structure and connectivity in species selected

1. Is there evidence of population genetic structure across the sampled range?

   H_0: No population structure - individuals at all sample sites analysed are in panmixia.

   H_{1A}: Weak population structure - suggestive of high gene flow and/or N_e among sample sites.

   H_{1B}: Regional population structure - suggestive of reduced gene flow among certain sample sites.

2. Is the British MPA network sufficient to maintain connectivity in this species?

   H_0: No - evidence suggests the network is insufficient to maintain connectivity.

   H_{1A}: Yes - evidence suggests the network is sufficient to maintain connectivity.

   H_{1B}: Yes - some evidence that the network could maintain connectivity between certain areas.

Chapter 2: Meta-analysis of northeast Atlantic marine taxa shows contrasting phylogeographic patterns following post-LGM expansions

---

This chapter is based on a paper published in the journal *PeerJ*. The reference is given below and the full paper is available in the Appendix.

Jenkins TL, Castilho R, Stevens JR (2018) Meta-analysis of northeast Atlantic marine taxa shows contrasting phylogeographic patterns following post-LGM expansions. *PeerJ* **6**, e5684.

2.0 Abstract

Comparative phylogeography enables the study of historical and evolutionary processes that have contributed to shaping patterns of contemporary genetic diversity across co-distributed species. In this study, we explored genetic structure and historical demography in a range of coastal marine species across the northeast Atlantic to assess whether there are commonalities in phylogeographic patterns across taxa and to evaluate whether the timings of population expansions were linked to the Last Glacial Maximum (LGM). A literature search was conducted using Web of Science. Search terms were chosen to maximise the inclusion of articles reporting on population structure and phylogeography from the northeast Atlantic; titles and abstracts were screened to identify suitable articles within the scope of this study. Given the proven utility of mtDNA in comparative phylogeography and the availability of these data in the public domain, a meta-analysis was conducted using published mtDNA gene sequences. A standardised methodology was implemented to ensure that the genealogy and demographic history of all mtDNA datasets were reanalysed in a consistent and directly comparable manner. Mitochondrial DNA datasets were built for 21 species. The meta-analysis revealed significant population differentiation in 16 species and four main types of haplotype network were found, with haplotypes in some species unique to specific geographical locations. A signal of rapid expansion was detected in 16 species, whereas five species showed evidence of a stable population size. Corrected mutation rates indicated that the majority of expansions were estimated to have occurred after the earliest estimate for the LGM ($\sim$26.5 Kyr), while few expansions were estimated to have pre-dated the LGM. This study suggests that post-LGM expansion appeared to be common in a range of marine taxa, supporting the concept of rapid expansions after the LGM as the ice sheets started to retreat. However, despite the commonality of expansion patterns in many of these taxa, phylogeographic patterns appear to differ in the species included in this study. This suggests that species-specific evolutionary processes, as well as historical events, have likely influenced the distribution of genetic diversity of marine taxa in the northeast Atlantic.

2.1 Introduction

Comparative phylogeographic studies present opportunities to explore how historical events may have helped shape patterns of genetic structure amongst co-distributed species (Avise et al. 1987; Avise 2009; Hickerson et al. 2010). Patterns of concordant phylogeographical structure across multiple taxa are

particularly informative because, while some patterns of spatial genetic structure may be caused by species-specific evolutionary processes, patterns common across multiple taxa may suggest similar evolutionary histories, such as common barriers to gene flow (Avise 2009; Hickerson et al. 2010). These findings can be important for conservation because of the potential to modify management actions in the light of the differing phylogeography of multiple species across the same geographical area (Pelc et al. 2009; Toonen et al. 2011; Heyden et al. 2014; Liggins et al. 2016). In marine biology, such comparative studies have made important contributions to our understanding of how historical events, such as the Pleistocene glaciations, have helped shape the spatial patterns of contemporary genetic diversity of marine taxa (Patarnello et al. 2007; Maggs et al. 2008; Marko et al. 2010; Ni et al. 2014).

The Pleistocene epoch was characterised by recurrent glaciations and intensive fluctuations in climate that periodically influenced the spatial distributions of plants and animals (Hewitt 1999; Hofreiter & Stewart 2009). The most recent glacial period began approximately 115 Ka and nearly all ice sheets were at their maximum (Last Glacial Maximum, LGM) between 26.5-19 Ka (Clark et al. 2009). The advances of the Northern Hemisphere ice sheets led to significant changes in temperature and sea levels (Lambeck & Chappell 2001). This must have had profound implications for habitat availability and the population persistence of coastal species-large parts of species' ranges would have been reduced, while other species may have survived in glacial refugia (Maggs et al. 2008; Provan & Bennett 2008). As the ice retreated and the sea level rose, a number of individuals from refugial populations may have dispersed and recolonised areas unavailable during the glaciation (Hewitt 2000). Changes in latitudinal ranges and population sizes can have distinct effects on the genetic architecture of a species due to the competing processes of mutation, drift and selection; moreover, the deep molecular divergence reported in taxa associated with several known European refugia suggests repeated expansion and contraction of conspecific populations were common throughout the Pleistocene (Hewitt 2004).

In the northeast Atlantic, the ice sheets extended as far south as Britain and Ireland, leaving an ice-free zone in mid-southern England, with possibly a small area in southwest Ireland free of ice (Chiverrell & Thomas 2010). However, the predicted extent of ice coverage across southern Ireland and the Celtic Sea differs among studies (e.g. Taberlet et al. 1998; Hughes et al. 2016). The advance of the ice sheets led to a drastic drop in sea levels in the English Channel, resulting in the complete emersion of the channel between England and France, except for a palaeo-river that extended across the continental margin (Ménot et al. 2006). This suggests that extant coastal communities inhabiting these areas are likely recolonisers originating

from glacial refugia. It has been suggested that Hurd Deep, a trench in the English Channel (Fig. 8), might have persisted as a marine lake during the LGM, thereby acting as a potential glacial refugium (Provan et al. 2005; Hoarau et al. 2007). Other areas further south, including Brittany (Coyer et al. 2003) and the Iberian Peninsula (Hoarau et al. 2007; Neiva et al. 2012) (Fig. 8), have also been postulated to act as refugia during the LGM. This was supported by high levels of genetic diversity found at these areas in the species studied, a key signature indicative of glacial refugia (Provan & Bennett 2008).

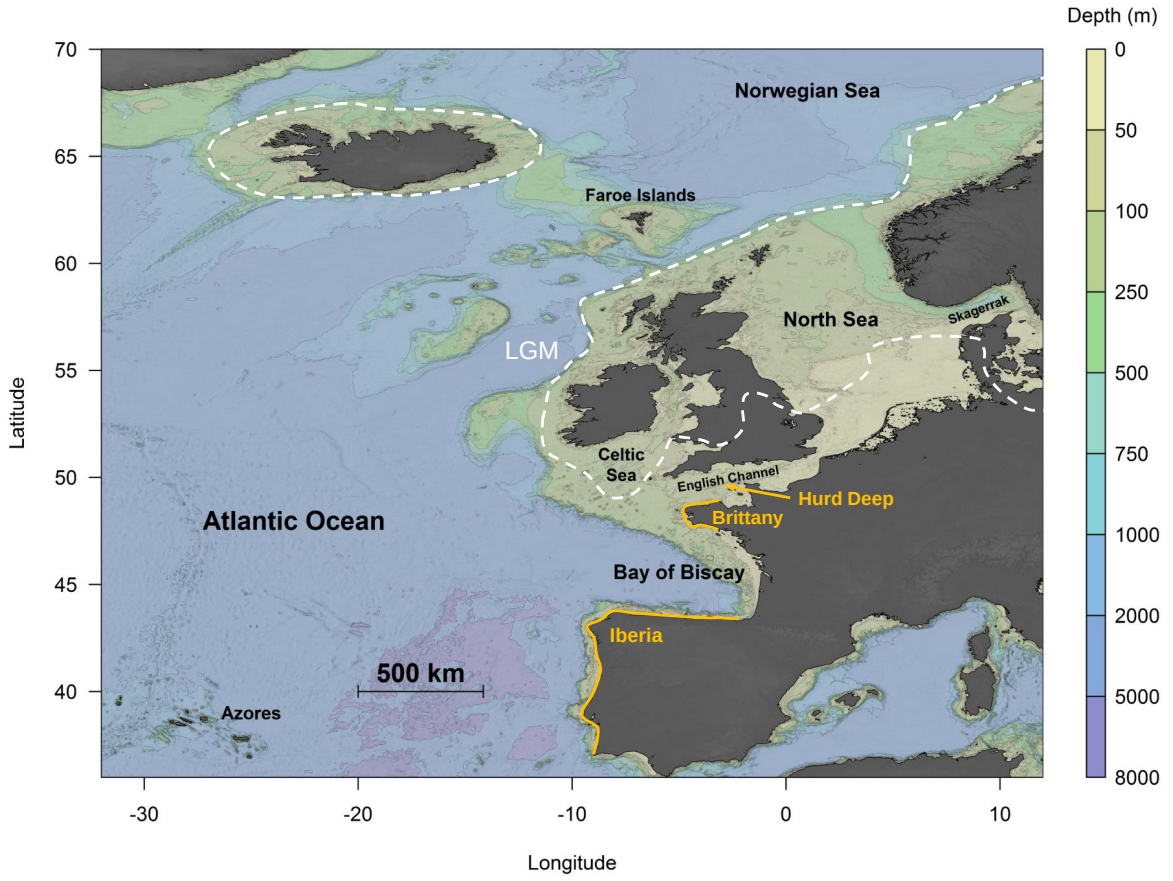


Figure 8: Topographical map of the northeast Atlantic Ocean. The white dotted lines represent the maximum extent of ice cover during the Last Glacial Maximum (LGM) (redrawn from Hughes et al. 2016). Orange lines indicate putative refugia: Hurd Deep, Brittany and Iberia.

Studies of single-species phylogeography across the northeast Atlantic are common; yet, because of the differences in molecular methodologies and analytical approaches, it can be difficult to compare results reliably. By applying a consistent methodology across all studies, this standardises the analysis (Harrison 2011), enabling patterns of phylogeography to be explored and compared within and across taxa. Two comparative meta-analyses in the Atlantic Ocean have been published to-date: the first explored the feasibility of distinguishing genetic

signatures of periglacial refugia from southern refugia in eight benthic marine species (Maggs et al. 2008), and the second looked for concordance among phylogeographical breaks around the southeast coast of the United States of America (Pelc et al. 2009). Systematic meta-analyses across diverse taxa in other seas and oceans have proved useful for exploring broad patterns of phylogeography (e.g. Patarnello et al. 2007; Kelly & Palumbi 2010; Marko et al. 2010; Ni et al. 2014); for example, one study of rocky-shore taxa from the northeastern Pacific found that 36 % of species showed evidence of population expansions associated with the LGM, while 50 % exhibited demographic patterns consistent with stable effective population sizes (Marko et al. 2010). However, such a study for marine taxa across the northeast Atlantic has yet to be undertaken.

In this study, we reanalyse available mitochondrial (mt)DNA data to compare the phylogeography of coastal benthic and demersal organisms across the northeast Atlantic (Fig. 8), an area characterised by complex oceanography and historical biogeographical events, such as the Pleistocene glaciations. Specifically, our aims were: (i) to identify commonalties (or otherwise) in contemporary genetic structure; (ii) to re-examine historical demography to test for signatures of population expansions; and (iii) to estimate the timings of any expansions detected. We discuss our findings in the context of the Pleistocene glaciations, asking in particular whether the LGM affected the phylogeography of marine taxa concordantly or discordantly.

2.2 Materials and methods

2.2.1 Literature search

To compare the phylogeography of benthic and demersal organisms across the northeast Atlantic, we undertook a meta-analysis of molecular phylogeographic studies. A literature search was conducted using Web of Science (Thomson Reuters) in February 2015. Search terms were chosen to maximise the inclusion of articles reporting on population structure and phylogeography from the northeast Atlantic. The following sets of Boolean search terms were submitted to the Advanced Search Tool: (1) gene flow OR population structure OR genetic diversity OR phylogeograph*; (2) marine OR intertidal OR subtidal OR estuar*; and (3) Atlantic. Titles and abstracts were screened to identify suitable articles within the scope of this study and only articles that matched the following criteria were retained: (a) organisms were fully marine or estuarine throughout their life history (diadromous species were excluded); (b) studies of temporal changes, hybridisation or introgression from closely related species were omitted; (c) the

study included at least three sampling sites from within the northeast Atlantic (Fig. 8 – sites outside of this area were not considered); (d) datasets contained a minimum of five individuals per site and a total sample size of at least 50; and (e) the study included latitude and longitude of the sampling sites or a detailed description or map which provided sufficient detail to determine the geographical location of sample origins. Given the proven utility of mtDNA in comparative phylogeography (e.g. Patarnello et al. 2007; Ni et al. 2014) and the availability of these data in the public domain, a meta-analysis was conducted using published mtDNA gene sequences.

2.2.2 Data reanalysis

A standardised methodology was implemented to ensure that all mtDNA datasets were reanalysed in a consistent and directly comparable manner. Data analyses in the original studies were far from consistent, particularly with respect to the analysis of haplotype networks and historical demography. The majority of studies reported information about population structure; however, in several instances the studies included additional samples outside of the northeast Atlantic in their analysis. Therefore, standardised tests of population structure were undertaken *de novo* for each species. Sites that were genetically homogeneous (as described by the original authors) and which were spatially close or situated in the same geographical region were combined in some datasets. This ensured that phylogeography within and across seas was examined in this meta-analysis. Population differentiation was examined using global values of Jost's D (Jost 2008) and F_{st} (Weir & Cockerham 1984) using the `fastDivPart` function from the R package diveRsity (Keenan et al. 2013; R Core Team 2016) and significance was assessed using 10,000 permutation replicates.

To examine the genealogical relationships within species, haplotype networks were constructed using the `haploNet` function from the R package pegas (Paradis 2010). Tajima's D (Tajima 1989), Fu's F_s (Fu 1997) and Ramos-Onsins' R_2 (Ramos-Onsins & Rozas 2002) neutrality tests were performed in DnaSP v5.10 (Librado & Rozas 2009) to determine whether each species carried a signal that deviated from neutrality (significance was assessed using 10,000 bootstrap replicates). Mismatch analyses (frequency of pairwise nucleotide-site differences between sequences) were carried out using the population growth-decline model in DnaSP to further examine the demographic history, and Harpending's raggedness index (r) (Harpending 1994) was used to evaluate the fit of the observed distribution to the growth-decline model (10,000 bootstrap replicates). A

non-significant index suggests that the observed data have a relatively good fit to the growth-decline model. In contrast, a significant index is indicative of a stable population which is typically thought to show a 'ragged', multi-modal mismatch (Harpending 1994).

The equation $t = \tau/(2\mu k)$ was used to estimate the timing of a population expansion (t), where τ is the date of the expansion measured in units of mutational time (Tau – estimated using DnaSP), μ is the mutation rate per site per year and k is the sequence length. In addition, Bayesian Skyline Plots (BSPs) were run using BEAST2 v2.5.0 (Drummond et al. 2005; Bouckaert et al. 2014). BEAST2 uses a Markov chain Monte Carlo (MCMC) sampling procedure to estimate N_e through time based on the temporal distribution of coalescences in gene genealogies. For each dataset, the substitution model was selected using bModelTest (Barido-Sottani et al. 2018), which uses reversible jump MCMC that allows the Markov chain to jump between states representing different possible substitution models. A strict clock and a coalescent Bayesian Skyline prior was implemented. Each run consisted of 100 million steps with a burn-in of one million and parameters were sampled every 10,000 steps. Chain convergence and BSPs were analysed with Tracer v1.7.1 (Rambaut et al. 2018).

Recent studies have shown that the use of mutation rates derived from ancient calibration dates or from phylogenetic analyses may not be appropriate for studies at the population level (Ho et al. 2008, 2011). In this study, therefore, mutation rates were chosen based on the most recent calibration date available for the closest taxonomic relative (Appendix A1). In published studies where a mutation rate was not specified, the genetic distance provided by the study was divided by the date of the calibration event (in Myr) to obtain a % mutation rate per Myr. For cases where only calibration dates older than 5 Myr were available for the species and gene of interest, a three-fold correction in mutation rate was applied to the original rate to control for the potential time-dependency of molecular rates. This adjustment was implemented because rates have been found to vary by three to six-fold for several marine species when calibration dates younger than 5 Myr vs. older dates have been tested (Crandall et al. 2012; Laakkonen et al. 2015). A range of mutation rates based on the rates reported by previous studies were used to calculate a minimum, maximum and average time estimate since a population expansion.

2.3 Results

2.3.1 Literature search

The initial search using Boolean terms identified 1,120 articles, which was reduced to 56 articles after the titles and abstracts were examined and the search criteria were applied. The final database for the meta-analysis consisted of mtDNA gene sequence data from 21 studies (Table 2); some studies from the previous step were not included due to the use of RFLPs in mtDNA or because some mtDNA datasets were not publicly available. The final database spanned several taxonomic groups, with fishes, molluscs and crustaceans accounting for the majority of species (81%). The most common mtDNA gene across all studies was cytochrome oxidase I (COI), followed by cytochrome b (Cyt *b*), the control region (CR) and the intergenic spacer region (IGS). COI was the most commonly used gene for invertebrate studies, IGS for macroalgae, and studies of fish used either the CR or the Cyt *b* gene.

2.3.2 Genetic structure

Sixteen species showed significant global Jost's D and F_{st} values, indicative of population differentiation (Table 3), while the remaining five species showed little evidence of population differentiation. Across the 21 datasets, four different types of haplotype network (Fig. 9) were putatively identified based on the structure of the networks:

(i) A 'Star' network (Fig. 9a), in which a single, widespread haplotype is typically positioned at the centre of the network and is thought to be the ancestral haplotype. Additional haplotypes are linked to this dominant haplotype by a single (or a few) mutational step(s), suggesting these haplotypes are the product of recent mutation events. Eight species showed this type of relationship (*Celleporella hyalina*, *Conger conger*, *Nassarius nitidus*, *Nassarius reticulatus*, *Palinurus elephas*, *Pelvetia canaliculata*, *Pomatoschistus microps* and *Raja clavata*). In one case, the dominant haplotype had far fewer connections than a low-frequency haplotype in the network, making it difficult to distinguish the centre of the network with confidence (*Pomatoschistus microps*);

(ii) A 'Complex star' network (Fig. 9b), in which there are multiple high-frequency haplotypes and connections. Six species showed this type of relationship (*Carcinus maenus*, *Cerastoderma edule*, *Maja brachydactyla*, *Pomatoschistus minutus*, *Solea solea*, *Symphodus melops*);

(iii) A 'Reciprocally monophyletic' network (Fig. 9c), in which more than one lineage is apparent and each lineage is linked by a long branch associated with

Table 2: List of the papers used in the meta-analysis and a summary of the information extracted from each study.

Taxon Species	mtDNA gene	No. sites; *N*	Sampling site distribution	Larval development	No. of lineages	Reference
Crustacean						
Carincus maenas	COI	13; 200	SW Spain to Norway	PLD, long	1	Roman & Palumbi (2004)
Maja brachydactyla	COI	13; 291	SW Spain to W Ireland	PLD, 2-3 wk	1	Sotelo et al. (2008)
Neomysis integer	COI	9; 379	SW Spain to E Scotland	No PLD, brooder	1	Remerie et al. (2009)
Palinurus elephas	COI	6; 119	SW Spain to W Scotland	PLD, up to 1 yr	1	Palero et al. (2008)
Fish						
Conger conger	CR	4; 232	Azores to Ireland	Leptocephalus, up to 2 yr	1	Correia et al. (2012)
Dicentrarchus labrax	CR	9; 93	Bay of Biscay to Norway	PLD, 8-12 wk	1	Coscia & Mariani (2011)
Labrus bergylta	CR	7; 279	W Ireland to Norway	PLD, 37-49 d	1	D'Arcy et al. (2013)
Pomatoschistus microps	Cyt *b*	10; 232	Bay of Biscay to Norway	PLD, 6-9 wk	1	Gysels et al. (2004)
Pomatoschistus minutus	Cyt *b*	8; 165	S Portgual to Norway	PLD, unknown	1	Larmuseau et al. (2009)
Raja clavata	Cyt *b*	9; 315	Azores to North Sea	No PLD, oviparous	1	Chevolot et al. (2006)
Solea solea	Cyt *b*	10; 645	Bay of Biscay to Skagerrak	PLD, up to 3 wk	1	Cuveliers et al. (2012)
Symphodus melops	CR	10; 263	S Portugal to Skagerrak	PLD, 14-25 d	1	Robalo et al. (2012)
Macroalgae						
Pelvetia canaliculata	IGS	15; 429	Portugal to Norway	External fertilisation	1	Neiva et al. (2014)
Mollusc						
Cerastoderma edule	COI	12; 300	Portugal to Norway	PLD, up to 4 wk	1	Krakau et al. (2012)
Macoma balthica	COI	15; 339	Bay of Biscay to North Sea	PLD, 2-5 wk	2	Becquet et al. (2012)
Modiolus modiolus	COI	4; 73	Irish Sea to Norway	PLD, up to 24 wk	2	Halanych et al. (2013)
Nassarius nitidus	COI	3; 62	NW Spain to Sweden	PLD, 4-8 wk	1	Couceiro et al. (2012)
Nassarius reticulatus	COI	6; 156	S Portugal to UK	PLD, 4-8 wk	1	Couceiro et al. (2007)
Polychaete						
Owenia fusiformis	COI	11; 283	Portugal to North Sea	PLD, up to 28 d	3	Jolly et al. (2006)
Pectinaria koreni	COI	10; 289	Portugal to North Sea	PLD, up to 15 d	2	Jolly et al. (2006)
Bryozoan						
Celleporella hyalina	COI	9; 63	NW Spain to Iceland	PLD, 1-4 h	1	Gómez et al. (2007)

MtDNA, mitochondrial DNA; No. of sites, number of sampling sites; *N*, number of individuals; PLD, pelagic larval duration.

Table 3: Summary statistics for each species. Population differentiation and demographic statistics are shown. In all statistical tests, significance was assessed using 10,000 permutations or bootstraps replicates.

Species	Differentiation		Demography				
	Jost's *D*	F_{st}	Tajima's *D*	F_s	R_2	*r*	Expansion
Crustacean							
Carcinus maenas	0.584***	0.157***	-1.73*	-40.36***	0.034*	0.018	Yes
Maja brachydactyla	0.298***	0.045***	-1.86**	-33.72***	0.028*	0.030	Yes
Neomysis integer	0.956***	0.554***	0.14	-0.954	0.024	0.086	No
Palinurus elephas	0.023	0.000	-2.31***	-30.19***	0.019*	0.094	Yes
Fish							
Conger conger	0.124	0.000	-2.58***	-211.1***	0.012***	0.031	Yes
Dicentrarchus labrax	0.540*	0.031*	-1.88**	-21.52***	0.047*	0.011	Yes
Labrus bergylta	0.672***	0.135***	-0.53	-49.35***	0.074	0.024	Yes
Pomatoschistus microps	0.391***	0.385***	-1.39	-17.90***	0.044	0.215	Yes
Pomatoschistus minutus	0.652***	0.100***	-1.96**	-90.56***	0.034*	0.015	Yes
Raja clavata	0.375***	0.330***	-0.09	-2.340	0.076	0.309	No
Solea solea	0.049	0.002	-2.02***	-131.9***	0.021**	0.221	Yes
Symphodus melops	0.578***	0.349***	-1.70*	-50.52***	0.032*	0.086	Yes
Macroalgae							
Pelvetia canaliculata	0.689***	0.482***	-1.53*	-19.02***	0.036	0.043	Yes
Mollusc							
Cerastoderma edule	0.662***	0.304***	-2.24***	-34.47***	0.019**	0.033	Yes
Macoma balthica	0.702***	0.470***	–	–	–	–	–
lineage 1	0.551***	0.434***	-0.80	-3.773	0.053	0.241	No
lineage 2	0.007	0.000	-0.99	-1.110	0.089	0.173	No
Modiolus modiolus[a]	0.083	<0.001	-1.79*	-11.91***	0.045*	0.156	Yes
Nassarius nitidus	0.222***	0.302***	-1.49*	0.028	0.049*	0.446	No
Nassarius reticulatus	0.047	0.000	-2.51***	-48.33***	0.016**	0.080	Yes
Polychaete							
Owenia fusiformis	0.788***	0.055***	–	–	–	–	–
lineage 1	0.636	0.001	-2.34***	-114.8***	0.024**	0.020	Yes
lineage 2	0.734	0.012	-2.06**	-55.00***	0.030**	0.008**	Yes
lineage 3	0.050	0.000	-1.26	-3.934**	0.084	0.080	Yes
Pectinaria koreni	0.596***	0.112***	–	–	–	–	–
lineage 1	0.638*	0.024**	-1.99**	-76.48***	0.027**	0.021	Yes
lineage 2	0.390	0.050	-2.63***	-54.02***	0.018***	0.029*	Yes
Bryozoan							
Celleporella hyalina	0.513***	0.488***	-1.35	-0.554	0.063	0.061	No

*<0.05, **<0.01, ***<0.001.
F_s, Fu's F_s; R_2, Ramos-Onsins' R_2; *r*, Harpending's raggedness index.
[a]Only statistics for lineage 1 are shown.

numerous mutations. Four species showed this type of relationship (*Macoma balthica*, *Modiolus modiolus*, *Owenia fusiformis* and *Pectinaria koreni*);

(iv) A 'Complex mutational' network (Fig. 9d), in which some branches were separated by a very large number of mutations, while other branches had contrarily one or two mutations. Three species showed this type of relationship (*Dicentrarchus labrax*, *Labrus bergylta* and *Neomysis integer*). In most cases, a dominant haplotype was present and was presumed to be the ancestral form. However, *Neomysis integer* presented an unusual network in which a distinct ancestral haplotype was not apparent and the centre of the haplotype network was not readily distinguishable. Haplotype networks for all species are available in the Appendix (A2).

2.3.3 Historical demography

Historical demography was inferred for each species based on the observed mismatch distribution, neutrality tests and the raggedness index (Table 3). Four main types of mismatch distributions were observed: unimodal, skewed unimodal, multimodal and bimodal (Fig. 10). Unimodal is associated with a sudden population expansion (e.g. *Maja brachydactyla*; Fig. 10a), and skewed unimodal is generally associated with a recent expansion or bottleneck (e.g. *Nassarius reticulatus*; Fig. 10b). Multimodal (e.g. *Labrus bergylta*; Fig. 10c) and bimodal (e.g. *Macoma balthica*; Fig. 10d) are usually associated with constant population size. However, previous research has suggested that bimodal peaks may indicate the presence of two distinct lineages (e.g. Alvarado-Bremer et al. 2005), which would potentially violate the assumptions of coalescent theory if analysed as one 'genetic' population. In this case, the first peak would represent intra-clade pairwise differences, whereas the second peak would likely represent more ancient inter-clade pairwise differences (Fig. 10D). For each instance of bimodality, the haplotype network was inspected for evidence of two or more lineages. The networks indicated that more than one distinct lineage was evident for all bimodal mismatches (*Macoma balthica*, *Modiolus modiolus*, *Owenia fusiformis* and *Pectinaria koreni*) and, therefore, mismatch analysis and neutrality tests were carried out on each lineage separately. These analyses were not conducted for lineage 2 of *Modiolus modiolus* due to the small number of individuals (N = 3) comprising this lineage. Mismatch distributions for all species are available in the Appendix (A3).

Neutrality statistics for testing the drift-mutation equilibrium (Tajima's D, F_s and R_2) were found to be contrasting between species (Table 3). These tests tended to

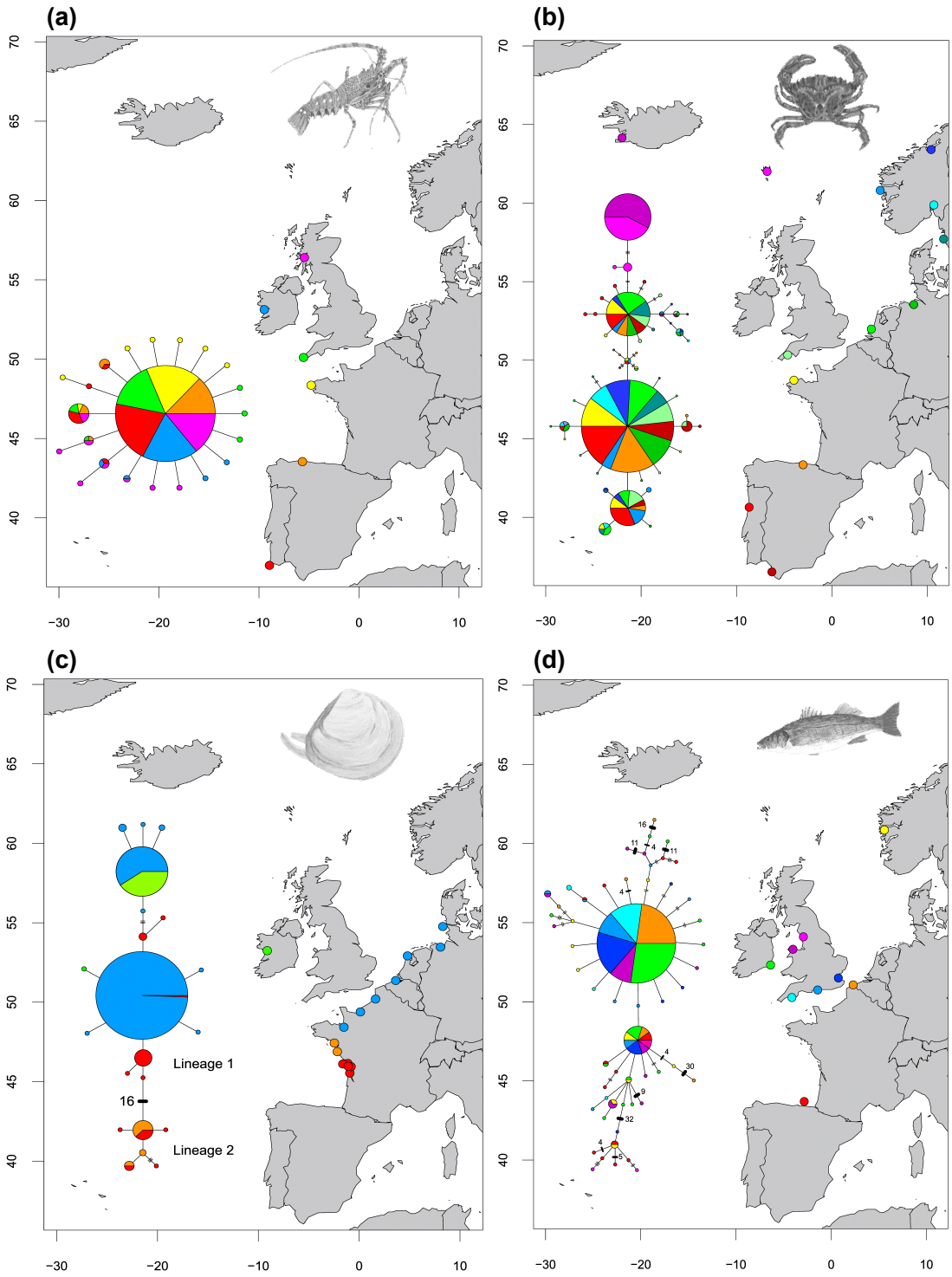


Figure 9: Haplotype networks showing four different network structures: (a) 'star' (*Palinurus elephas*), (b) 'complex star' (*Carcinus maenas*), (c) 'reciprocally monophyletic' (*Macoma balthica*) and (d) 'complex mutational' (*Dicentrarchus labrax*). Each circle represents a unique haplotype and the sizes of the circles are proportional to the haplotype frequencies for each network but are not comparable across studies. Each line represents one mutation step and two or more steps are indicated by bars or numbers. Colours inside the circles correspond to sites which have individuals represented in that particular haplotype.

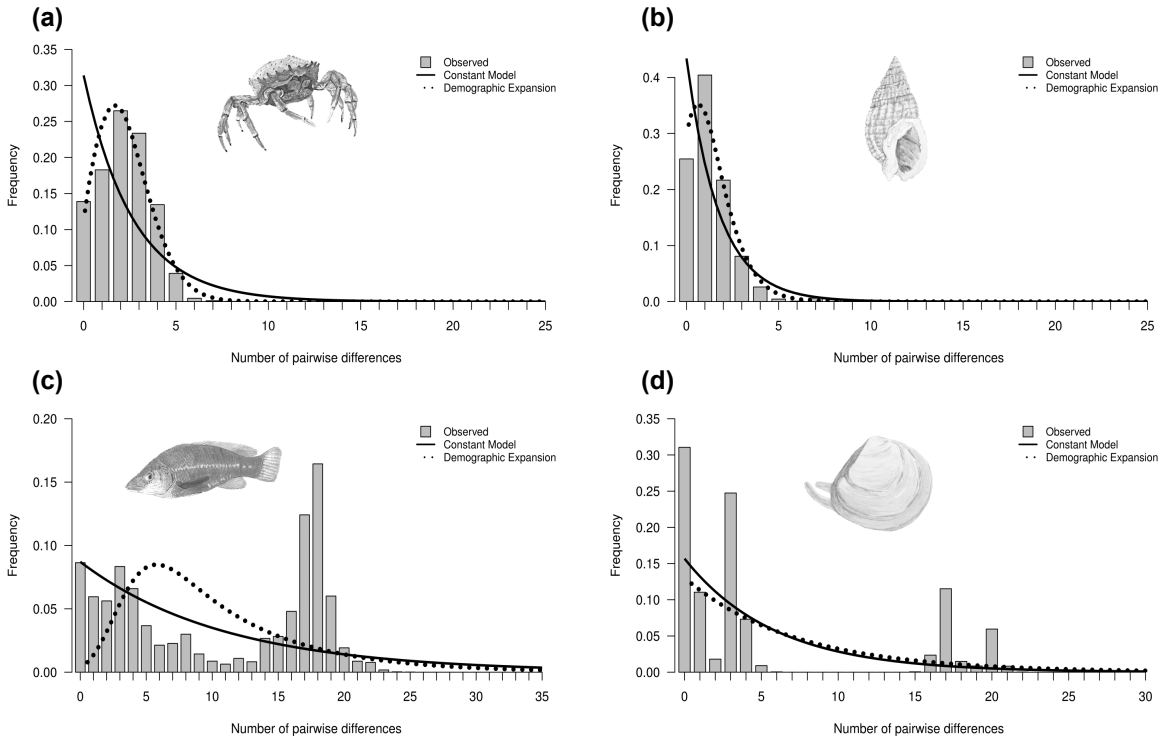


Figure 10: Mismatch distributions showing four different distributions: (a) unimodal (*Maja brachydactyla*), (b) skewed unimodal (*Nassarius reticulatus*), (c) multimodal (*Labrus bergylta*) and (d) bimodal (*Macoma balthica*). Unimodal and skewed unimodal distributions are generally associated with a sudden expansion and a recent sudden expansion, respectively. Multimodal and bimodal are thought to be associated with a constant population size (but see text). Bars represent the frequency of pairwise nucleotide differences between individuals. Curves correspond to the expected distribution fitted to the data under a model of constant population size (solid line) or demographic expansion (dotted line).

be significant for species that showed a star-shaped network and for which the mismatch graph was unimodal or skewed unimodal. This supported evidence that a signal of rapid population expansion was detected; however, a selective sweep can also produce the same genetic signal. Harpending's *r* suggested that two datasets departed from a model of demographic expansion (Table 3), but inspection of the mismatch graphs and neutrality tests indicated there was strong evidence to support a rapid population expansion (or selective sweep) in both datasets. No signatures of rapid population expansion were detected in five species (*Celleporella hyalina*, *Macoma balthica*, *Nassarius nitidus*, *Neomysis integer* and *Raja clavata*), suggesting a stable constant population size.

For the remaining 19 datasets (16 species, 19 including lineages), a historic population expansion was assumed and the timing of the expansion was estimated (Fig. 11). All expansions were found to take place during the Pleistocene or the Holocene epoch. Estimated timings for 17 datasets were after or overlapped the

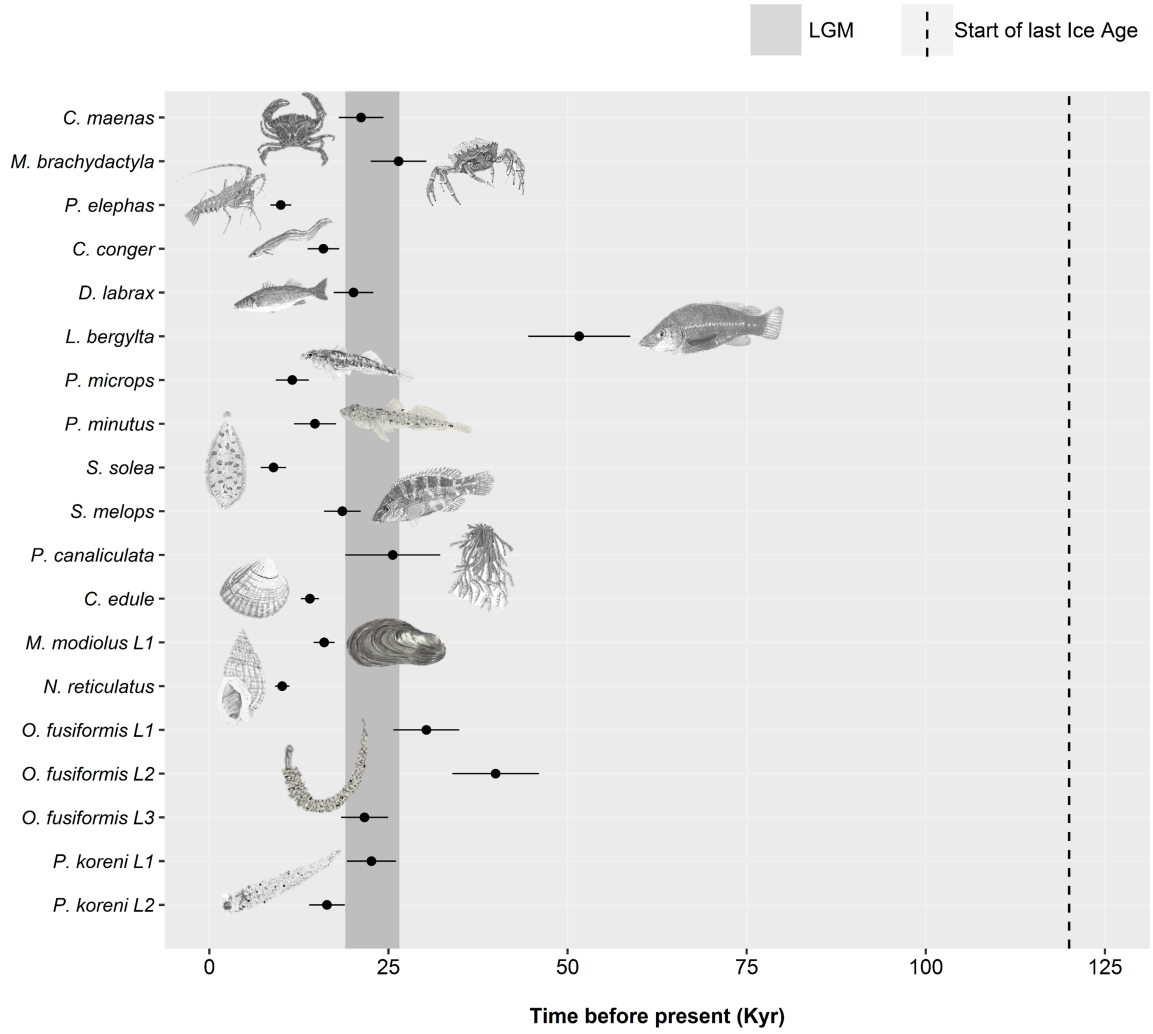


Figure 11: Estimated dates of expansion for species or lineages (L) in which the demographic expansion hypothesis was not rejected. A minimum and maximum time since expansion is plotted as horizontal bars for some datasets, estimated from a minimum and maximum mutation rate (Appendix A1). The beginning of the last glacial period (dotted line) and the estimated time-frame of the Last Glacial Maximum (grey shaded area) are displayed. Species are organised by taxa: crustaceans, *Carcinus maenas* - *Palinurus elephas*); fish, *Conger conger* - *Symphodus melops*; macroalgae, *Pelvetia canaliculata*; molluscs, *Cerastoderma edule* - *Nassarius reticulatus*; polychaetes, *Owenia fusiformis* - *Pectinaria koreni*.

earliest estimate for the LGM (∼26.5 Ka). Expansion estimates for one fish (*Labrus bergylta*) and one lineage of the polychaete *Owenia fusiformis* pre-dated the LGM but were still positioned during the last glacial period. Bayesian Skyline Plots (Fig. 12) were generally consistent with the results from the mismatch analyses. Among the 17 datasets for which the mismatch analyses estimated the time of an expansion to have occurred after the LGM, a rise in N_e post-LGM was apparent in 15 of these datasets, but the strength of the increase varied across datasets. In comparison to the mismatch analysis, the BSP for *L. bergylta* (Fig. 12f) and *O.*

fusiformis lineage 2 (Fig. 12p) indicated a population expansion after the earliest estimate for the LGM as opposed to pre-dating the LGM. In addition, although the mismatch analyses inferred a post-LGM expansion for *M. modiolus* lineage 1 (Fig. 12m) and *O. fusiformis* lineage 3 (Fig. 12q), BSPs generally suggested N_e was constant after the LGM.

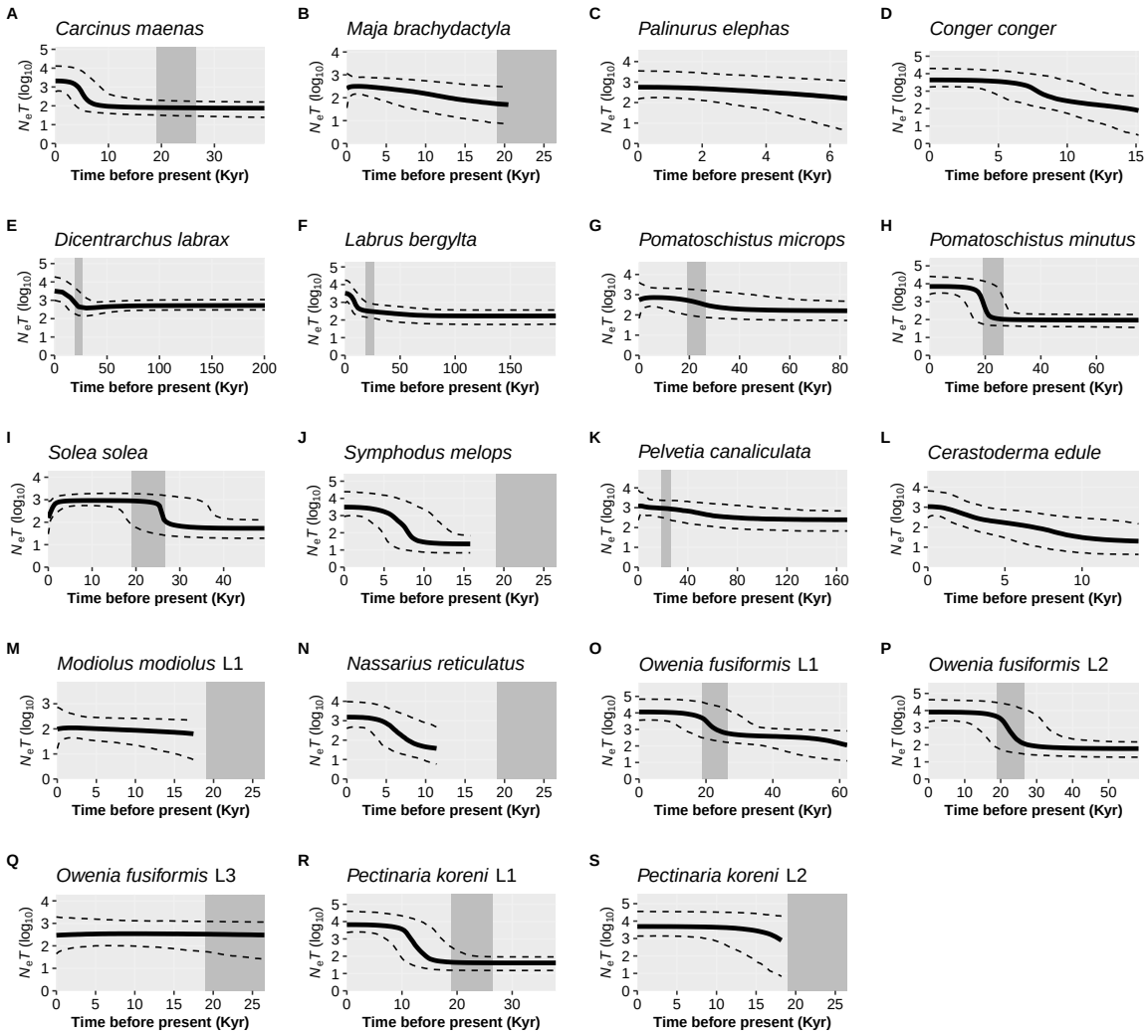


Figure 12: Bayesian Skyline Plots for species or lineages (L) in which the demographic expansion hypothesis was not rejected. Solid black lines show the median effective population size over time (N_e = effective population size and T = generation time); dashed black lines represent the 95% confidence intervals. The estimated time-frame of the Last Glacial Maximum is denoted by the area shaded dark grey. Species are organised by taxa: crustaceans, *Carcinus maenas* (a), *Maja brachydactyla* (b), *Palinurus elephas* (c); fish, *Conger conger* (d), *Dicentrarchus labrax* (e), *Labrus bergylta* (f), *Pomatoschistus microps* (g), *P. minutus* (h), *Solea solea* (i), *Symphodus melops* (j); macroalgae, *Pelvetia canaliculata* (k); molluscs, *Cerastoderma edule* (l), *Modiolus modiolus* lineage 1 (m), *Nassarius reticulatus* (n); polychaetes, *Owenia fusiformis* lineage 1 (o), *O. fusiformis* lineage 2 (p), *O. fusiformis* lineage 3 (q), *Pectinaria koreni* lineage 1 (r), *P. koreni* lineage 2 (s).

2.4 Discussion

The results of this study show a range of contemporary genetic patterns across the coastal marine taxa analysed in the northeast Atlantic. In general, genealogical patterns were not uniform within taxonomic groups, though common patterns were observed in both polychaete species, which implies that historical events may have affected these polychaete species similarly. Most species (76 %) showed evidence of population structuring, suggestive of restricted contemporary or historical gene flow between the sites studied. Of the species that exhibited no population differentiation, all five species have a pelagic larval phase, with a pelagic larval duration (PLD) ranging from up to three weeks (*S. solea*) to a year or more (*P. elephas* and *C. conger*) (Table 2). However, most of the species that demonstrated significant population differentiation also had a pelagic larval phase, ranging from a relatively short PLD of 1-4 h (*C. hyalina*) to a relatively long PLD of 8-12 weeks (*D. labrax*) (Table 2). Although speculative, taken altogether, this may suggest that larval development and PLD could be important factors in maintaining gene flow in some, but not all, of these species; however, more evidence is needed to confirm this. Indeed, whether a general correlation exists between PLD and genetic differentiation measures remains unclear because some studies have reported poor correlations between the two (Weersing & Toonen 2009; Kelly & Palumbi 2010; Riginos et al. 2011), while other studies have reported the opposite (Siegel et al. 2003; Selkoe & Toonen 2011) suggesting that PLD and genetic metrics can indeed reflect scales of dispersal if the sampling design is robust (Selkoe & Toonen 2011). Other factors could also explain a lack of correlation between PLD and genetic differentiation such as: (i) temporal and spatial fluctuations in PLD within a species, (ii) habitat availability and suitability, and (iii) delayed metamorphosis (Pechenik 1990, 2006). As a result, speculative relationships between PLD and genetic differentiation should be interpreted with caution.

In some of the species studied, certain geographical areas were dominated by a particular haplotype that was rarely or not present in other areas across the sampled range. For example, the green crab *Carcinus maenas* showed highly significant differentiation and distinctive haplotypes in the Faroe Islands and Iceland, a pattern detected by the original authors who subsequently concluded that a deep-water barrier to dispersal in green crabs was the driver of this pattern (Roman & Palumbi 2004). A similar pattern was also observed for two species around western Ireland in the northeast Atlantic. In *Celleporella hyalina* and *Macoma balthica*, distinct haplotypes composed a population around western Ireland; however, unique haplotypes were not apparent in other species analysed

in this study with similar sampling coverage (e.g. *Labrus bergylta*, *Palinurus elephas* and *Pelvetia canaliculata*). A discrepancy in genetic structure between species at this spatial scale has also been observed between two temperate octocoral species (*Eunicella verrucosa* and *Alcyonium digitatum*) using microsatellite markers, whereby northwest Ireland samples were found to be genetically isolated from other northeast Atlantic samples in *E. verrucosa*, but not in *A. digitatum* (Holland et al. 2017). This suggests that historical or contemporary gene flow between areas in the northeast Atlantic and western Ireland is likely possible, but in some cases the spatial patterns of genetic structure could be influenced by other processes such as strong selection pressures, species-specific life history traits, demographic fluctuations, or range expansions occurring at different times in different species (Hellberg 2009).

2.4.1 Demographic history

Demographic history was variable across species in the northeast Atlantic, as evidenced by both the diverse structuring of the haplotype networks and the observed mismatch distributions within species. The presence of one or more lineages and the complexity of mutational patterns in several networks suggested some species have undergone pronounced changes in their demography and genealogy. Connections with large mutation steps separating some haplotypes are indicative of deep phylogenetic splits in the genealogies and suggests the persistence of old populations in these species. Accumulating new mutations is a relatively slow process and, therefore, sufficient time since coalescence must have elapsed to facilitate these large sequence divergences (Avise 2009).

In the northeast Atlantic, the LGM has often been viewed as a possible explanation for discrepancies in genealogies and for rapid population expansions via recolonisation as glaciers started to retreat from their maximum positions (Hewitt 2004). In this study, we detected rapid expansions in many different taxa, of which the majority were estimated to occur after the LGM. This supports evidence for post-LGM expansions, possibly from periglacial refugia (Maggs et al. 2008) or via recolonisation of areas previously affected by the Northern Hemisphere ice sheets. These results are in contrast to the northeast Pacific where regional persistence during the LGM appeared to be common in rocky-shore organisms (Marko et al. 2010). The conclusions of several previous studies reanalysed in this meta-analysis also detected rapid expansions (e.g. Jolly et al. 2006; Sotelo et al. 2008; Larmuseau et al. 2009); however, the authors of these studies estimated the dates of these expansions to have occurred pre-LGM. This discrepancy could be

due to the differences in mutation rates, whereby the original authors typically used rates derived from ancient calibrations, while in this study we attempted to use more recent calibration dates to correct for the potential time-dependency of molecular rates (Ho et al. 2011).

Of course, we acknowledge that the signal of deviation from neutrality we detected may, in some cases, be the result of a selective sweep and not a rapid expansion. This signal could be distinguished by incorporating multi-locus data; nevertheless, given that a variety of species in this study showed similar genealogical patterns consistent with demographic expansion, it seems likely that most of them did indeed experience demographic changes associated with the end of the LGM, rather than selective sweeps. Moreover, distinctive haplotypes were found in several species networks (*Pelvetia canaliculata*, *Pomatoschistus minutus*, *Owenia fusiformis* and *Pectinaria koreni*) to the south of where the Eurasian ice sheet is proposed to have extended during the LGM (Fig. 8). This finding suggests populations of these species may have survived in southern glacial refugia; though, as pointed out by some of the original authors, deep sequence divergences in some species (e.g. *O. fusiformis* and *P. koreni*) and the lack of a species-specific molecular clock calibration makes inferences about refugia challenging (Jolly et al. 2005, 2006).

It is difficult to suggest an explanation for the two expansions estimated to have pre-dated the LGM (using mismatch analysis), but which fall within the last glacial period. This pattern of pre-LGM expansion has also been reported in a number of previous studies for a variety of marine taxa (e.g. Hoarau et al. 2007; Marko et al. 2010; Ni et al. 2014; Almada et al. 2017). One potential explanation for this pattern is that sea level during the last glacial cycle did not decrease uniformly towards the level observed at the LGM, but oscillated rapidly over a period of 60 Ka to 30 Ka (see Fig. 3a in Lambeck et al. 2002). Therefore, it may be possible that we are detecting the signature of a population expansion during one of these sudden increases in sea level during the last glacial period. Alternatively, as the BSP analysis inferred a post-LGM expansion for these two datasets, this could be a limitation associated with the mismatch analysis approach, which does not consider genealogy, and may, therefore, produce a less precise estimation. In addition, the sample of genetic diversity for this species may not be representative (Karl et al. 2012) or the genetic signal we detected may have been the result of a selective sweep and not a rapid expansion.

The use of single marker mtDNA genealogies and coalescence theory can introduce challenges associated with the interpretation of data and these limitations should be acknowledged (Karl et al. 2012). For example, the

populations under study may have experienced multiple episodes of growth and decline; however, only the most recent expansion event can be detected using coalescence analysis and, in some cases, these events may not be sufficiently severe to be detected (Karl et al. 2012). In addition, coalescent histories can differ amongst loci because they can experience mutation and drift independently. Therefore, analysis of a single gene only gives insight into the coalescent history of that locus, which may not always be representative of population history. Analysis of multiple loci and genomics would help to alleviate these concerns, and would likely provide enhanced resolution for exploring the phylogeography of northeast Atlantic marine fauna.

Although population expansions were detected in a number of species in this study and also in the wider literature, populations of other marine species, including five from this study, have been found to remain stable throughout the LGM. As previously reported, not all coastal marine taxa appear prone to demographic changes during or after ice ages (Janko et al. 2007; Marko et al. 2010; Olsen et al. 2010). It is also important to acknowledge that earlier events in the Pleistocene and more ancient events that pre-date the Pleistocene may have helped shape the contemporary patterns of genealogical structure observed in this study.

2.4.2 Implications for conservation

Conservation and management plans are crucial for mitigating the loss of biodiversity from the effects of climate change and anthropogenic stressors in the marine environment. However, conservation plans have typically prioritised protecting contemporary patterns of biodiversity and seldom considered the evolutionary processes that generated them (Wright et al. 2015). As a result, genetic data have rarely been used to inform management, despite their relevance in conservation for describing/inferring patterns of genetic diversity, dispersal ability and evolutionary history, and for delineating cryptic species and evolutionary significant units (Beger et al. 2014).

Identifying discrete populations with high genetic diversity and detecting phylogeographic breaks to gene flow is of high importance for managers of MPA networks and conservation (Allendorf et al. 2010). Summarising broad patterns of genetic diversity and population structure across a broad range of marine species can facilitate the detection of major genetic breaks/barriers to gene flow and identify genetic diversity unique to specific areas. Such genetic diversity was found in several species analysed in this study, particularly in Iberian and western Ireland sites, and implies that marine conservation in the northeast Atlantic should attempt

to address these areas of unique genetic diversity which may be species-specific. This supports previous suggestions that managers should attempt to consider data from multiple species to position MPAs when trying to maximise ecological connectivity in their network (Marti-Puig et al. 2013; Pascual et al. 2017).

Although the results of this study are based on data from a single mtDNA gene and 21 species (albeit across a range of phyla), the patterns detected and discussed throughout this study may be informative as they represent the contemporary genealogical spatial structure of the matrilineal lineages and provide information about the evolutionary histories of these species and about potential cryptic species. While challenges exist in the interpretation of these data, conducting a meta-analysis of comparative phylogeography using multiple species can identify common signals of phylogeography, which can give clues about how populations of benthic marine species may respond to future changes in the environment.

2.4.3 Conclusions

The findings of this meta-analysis indicate that species in the northeast Atlantic do not show a uniform pattern of phylogeography, but rather a mixture of complex contemporary genealogical structure. Reanalysis of demographic histories indicated that a large proportion of the species included in this study have experienced post-LGM expansions, supporting the general expectation that rapid population expansions occurred after the LGM as the ice sheets started to retreat (Hewitt 2000, 2004). This suggests that regional extirpation during the LGM appears to be a common biogeographic history for many northeast Atlantic marine taxa. However, improvements in mutation rate estimates, as well as the incorporation of multi-locus markers and genomics, would likely provide greater accuracy and resolution for overcoming the challenges associated with single mtDNA genealogies, and for improving our understanding of phylogeography in the northeast Atlantic Ocean.

Chapter 3: Selecting taxa to study genetic connectivity between Marine Protected Areas

---

This chapter is based on a paper published in the journal *Marine Policy*. The reference is given below and the full paper is available in the Appendix.

Jenkins TL, Stevens JR (2018) Assessing connectivity between MPAs: selecting taxa and translating genetic data to inform policy. *Marine Policy* **94**, 165-173.

3.0 Abstract

Connectivity is frequently cited as a vital component of Marine Protected Area (MPA) networks and was formally identified as one of five key principles for marine network design in European waters. Yet, without the ability to demonstrate connectivity, it is impossible to be certain that sites designated within a MPA network do in fact constitute a network, when they may -irrespective of the diversity and rarity of the taxa within them- be in reality a set of unlinked habitats and associated species assemblages. However, the process of assessing connectivity between MPAs, and which taxa to include in assessments of connectivity, is often difficult and can be dependent on a variety of factors that may be beyond the control of managers, stakeholders and policymakers. In this chapter, a set of biological and methodological factors are highlighted, consideration of which may help to inform the selection of species for assessments of genetic connectivity between MPAs in a network. After cogitating these factors, two benthic species with differing life histories were selected as candidates for assessing connectivity between UK MPAs: the pink sea fan (*Eunicella verrucosa*) and the European lobster (*Homarus gammarus*). Exploring the population structure and genetic connectivity in these species will provide an empirical assessment of connectivity between MPAs; in addition, due to the important ecological and economical status of these two species, the results have the potential to inform the relevant conservation and fisheries management bodies.

3.1 Introduction

Connectivity is identified as a key component in the design of European Marine Protected Area (MPA) networks (OSPAR Commission 2013). However, changes to the definition of connectivity outlined in many different reports (OSPAR Commission 2010, 2013; Carr et al. 2014) suggest there is potential confusion or conflict amongst stakeholders and scientists concerning the exact definition and function of connectivity in the context of MPA networks. The most simplistic definition is taken from Palumbi (2003) whereby "*connectivity is the extent to which populations in different parts of a species range are linked by the movement of eggs, larvae or other propagules, juveniles or adults*" (OSPAR Commission 2013). In contrast, other reports have outlined a more detailed definition such that maintaining connectivity involves creating "...*ecologically connected and functional networks with 'corridors' or 'stepping stones' that facilitate the range shifts of populations and the movements of individuals and genes in response to ocean climate change*" (OSPAR Commission 2010), or that "...*the MPA network is well*

distributed in space and takes into account the linkages between marine ecosystems" (Carr et al. 2014).

Knowledge of connectivity is fundamental for optimising the location and size of MPAs to create a well-connected network (instead of individual unrelated MPAs) (Jones et al. 2007; Almany et al. 2009), and for evaluating the impacts of exploitation on the population dynamics of commercial marine species (Bernatchez et al. 2017). To understand connectivity, an ideal scenario might incorporate multiple sources of data informing on connectivity from many types of taxa within the boundaries of an MPA network; however, this is often impossible due to financial and logistical constraints. Instead, managers of MPAs have typically concentrated their efforts on species that are endangered or rare, and which may be on the brink of extirpation in parts of their range, or on so-called 'umbrella', 'keystone' or 'flagship' species (Simberloff 1998; Kalinkat et al. 2017). The concept of an umbrella species, a species whose protection indirectly protects many other species in an ecological community, is generally recognised as appealing for assessing connectivity. This is because the establishment of a network based on such data may extrapolate the benefits of preserving the connectivity of one focal species to other species in a community with similar life histories and dispersal traits. Hypothetically, a species associated with all three concepts (umbrella, keystone or flagship) would likely be the 'holy grail' species for studying connectivity between MPAs, but identification of such species (if indeed they exist) has continued to elude those involved in marine conservation. Moreover, for a variety of reasons (Table 4), the study of species that come close to satisfying the criteria of a 'holy grail' species may not be feasible and, therefore, compromises are needed to facilitate the collection of data that are informative about connectivity in a given system.

In this chapter, a number of biological and methodological factors are highlighted that should be considered before selecting taxa to assess genetic connectivity between MPAs. Subsequently, based on the set of factors highlighted, and existing knowledge of northeast Atlantic coastal taxa, two species are selected to explore spatial genetic structure and connectivity between British MPAs and across the wider northeast Atlantic.

3.2 Selecting taxa

The selection of appropriate taxa to use as surrogates for assessments of genetic connectivity between MPAs has seldom been discussed in the literature (but see Marti-Puig et al. 2013). Coastal benthic marine invertebrates are often good

candidates because they can be relatively abundant with large ranges, and dispersal is typically defined during a pelagic phase undertaken by an early life stage (e.g. eggs or larvae), while the adults remain relatively sedentary (Cowen & Sponaugle 2009). This type of development means connectivity is mainly dependent on local hydrological conditions (as well as species-specific traits) and, therefore, better reflects natural patterns of connectivity, as opposed to studying connectivity driven by organismal behaviour in motile and migratory species. Since patterns of genetic connectivity can vary between species over similar geographical areas (Coleman et al. 2011; Holland et al. 2017), it is also important to consider assessing connectivity in more than one species with differing biology/ecology. This allows species-specific genetic connectivity and patterns of connectivity common across taxa to be examined (Cowen et al. 2007; Marti-Puig et al. 2013).

3.2.1 Biological factors

Some biological features of candidate species can inevitably enhance the public appeal and societal impact of a study, while other features can limit the collection of samples and the interpretation of data generated by genetic markers (Table 4). For the purpose of promoting marine conservation, charismatic megafauna such as marine mammals and sharks frequently dominate awareness campaigns (flagship species) because they can raise funds and change public opinions and behaviour. Although many of these species may not be the best candidates for assessing MPA connectivity, these enigmatic animals are typically well-known by the wider public and benefit from a greater awareness and potential impact than other marine fauna. As a result, if a candidate species is poorly known to the public community, highlighting its importance for the conservation of an associated enigmatic species may have an equivalent effect (e.g. the interactions between kelp forests and sea otters, Lubchenco 2016).

Benthic marine invertebrates are generally not flagship species (but there are exceptions, e.g. pink sea fans). However, it is recognised that many benthic invertebrates have a crucial ecological role (e.g. mussel beds as ecosystem engineers / habitat builders) or are commercially exploited (e.g. scallops and lobsters), meaning they are either fundamentally important to the ecosystem or the local/regional economy, or both. This may encourage relevant management bodies and/or stakeholders to collaborate, to contribute funding and/or to share equipment (depending on the organisation's interests and capacity), all of which can serve to advance a particular project. For example, lobster fishermen have access to a potential myriad of individuals from which tissue samples can be obtained. Forming

Table 4: Biological and methodological factors to consider before selecting a species to assess genetic connectivity between Marine Protected Areas.

Factor	Description	Example	Significance
Biological			
Ecological importance	Does the organism have a fundamental importance to a functioning ecosystem?	Ecosystem engineers (e.g. mussel beds).	These species may be protected under legislation. Higher potential impact.
Flagship species	Is the species charismatic, well known to the public, and a poster child of conservation campaigns?	Many large megafauna including cetaceans and sharks. Some threatened invertebrates.	Greater public awareness/interest. Higher potential impact.
Economic importance	Is the species commercially exploited?	Fish and invertebrate coastal fisheries.	Opportunities for collaboration during sample collection.
Taxonomy	Is the taxonomy not well resolved and the organism hard to identify?	Sister species with very similar morphology. Very small organisms.	Species difficult to ID morphologically demand more resources and time. In some cases, a taxonomist or DNA barcoding method may be required for validation.
Biology and habitat knowledge	A sound knowledge of the biology and ecology of the organism.	Habitat / distribution, larval development, dispersal, etc.	Improve the testing of hypotheses. Improve interpretation of the results.
Methodological			
Sample collection	Collecting tissue samples from the organism for genetic analysis.	Feasibility / cost of collecting samples.	Protected species may require permits for tissue removal. Logistical barriers may limit sample collection in some areas (e.g. deep sea). Non-destructive tissue sampling advantagous for endangered or rare species.
Sample sites	The number of sampling sites and the spatial separation between sites.	Consider the number of sites needed in and around MPAs.	Adequate sampling sites in and out of MPAs could enhance hypothesis testing.
Sample sizes	The number of individuals per sampling site.	Consider the number of individuals needed to draw robust conclusions.	The number of individuals can be influenced by the choice of genetic marker. During data analysis, the power of the markers and sample sizes can be tested using various software.
DNA extraction	Extracting genomic DNA for analysis.	Consider tissue type and extraction protocol before sampling.	High quantity and quality DNA can be difficult to extract from some organisms and tissue types (e.g. crustacean exoskeleton) using standard kits.
Choice of genetic marker	Choosing a genetic marker that is polymorphic enough to investigate genetic patterns.	Microsatellites / SNPs.	The choice and number of markers will depend on the power and resolution required.
Availability of genetic markers	Are panels of markers already available for the organism?	Microsatellite / SNP panels.	This would avoid the need to develop markers *de novo*.

Table 5: Summary of the factors considered from the framework outlined in Table 4 for the pink sea fan (*Eunicella verrucosa*) and the European lobster (*Homarus gammarus*).

Factor	Pink sea fan	European lobster
Biological		
Ecological importance	IUCN Red Listed species and is specifically protected in English and Welsh waters. Is considered a flagship species in UK conservation.	Is a charismatic crustacean in the UK.
Flagship species	" "	" "
Economic importance	Not commericially exploited.	Active fishery that supports many coastal communities.
Taxonomy	Well resolved.	Well resolved.
Biology and habitat knowledge	Distribution and habitat well known.	Distribution, habitat, larval development and reproduction well known.
Methodological		
Sample collection	Samples in sufficient quantity have already been collected in a previous study.	Collaborations with fishermen and other stakeholders can facilitate the collection of a large number of samples.
Sample sites	Present in habitats within and outside of Marine Protected Area boundaries.	Present in habitats within and outside of Marine Protected Area boundaries.
Sample sizes	Samples in sufficient quantity have already been collected in a previous study.	Collaborations with fishermen and other stakeholders can facilitate the collection of a large number of samples.
DNA extraction	Protocol developed to extract high quality genomic DNA.	Protocol developed to extract high quality genomic DNA.
Choice of genetic marker	To explore fine-scale genetic patterns, more powerful markers may need to be developed anew.	Previous molecular work exists but limitations with resolution and sampling suggests markers will need to be developed anew.
Availability of genetics markers	Microsatellite markers have been developed and genotyped.	" "

these types of collaborations can facilitate access to a virtually unlimited number of samples depending on the fishery status, thereby avoiding the need to arrange dedicated sampling trips, and the associated costs and researcher time typically required for collection. Moreover, maintaining dialogue with such a stakeholder(s) may promote more effective communication of the potential benefits of the research and, ultimately, dissemination of the results (Shafer et al. 2015; Britt et al. 2018).

Other factors to consider include whether the biology and ecology of the candidate species is well known. This process starts, perhaps obviously, by accurate identification of the candidate species and avoiding the erroneous inclusion of closely related or cryptic species, which can drastically influence the results of population genetic structure analyses (Pante et al. 2015b). The difficulty of accurate taxonomic identification can be further exacerbated when the organism is very small; in some cases, a second opinion from an experienced taxonomist or molecular verification (e.g. DNA barcoding) may be required. In addition, a thorough understanding of the dispersal, life history and habitat of the candidate species will usually help to explain some of the genetic patterns observed, thereby improving interpretation of the genetic data.

3.2.2 Methodological factors

The sampling design of a study should be carefully considered prior to sample collection to ensure that the resulting genetic data are robust and applicable for assessments of genetic connectivity. This typically includes assessing whether the desired sampling strategy is feasible and that sufficient tissue samples from a broad enough range of sites can be taken for meaningful genetic analysis. For example, as suggested previously, if an organism is commercially fished, it may be possible to have tissue samples collected *in situ* by fishery personnel. Moreover, ensuring that samples of a species of interest are collected from both within the boundaries of a MPA network and from sites outside ensures that hypotheses about connectivity beyond MPA boundaries can be tested. This approach has provided useful data in several previous studies (Huserbraten et al. 2013; Puckett et al. 2014; Holland et al. 2017), allowing the performance of a MPA network to be evaluated for the species being studied.

Other factors to consider include the type(s) of tissue to sample and which genetic markers to use in assessments of population genetic structure. This is of critical importance because the type of tissue can profoundly influence the quantity and quality of DNA obtained post-extraction. For example, crustacean exoskeletal tissues, such as pleopods, are advantageous because they are easily obtained and

constitute a non-destructive tissue sample; however, extracting sufficient amounts of pure (contaminant-free) DNA from these tissue types can be extremely difficult using both conventional and kit-based protocols (Li et al. 2011). Moreover, obtaining high molecular weight, non-degraded DNA can be important for methods that utilise next-generation sequencing technology, for example, whole-genome sequencing and SNP discovery from RADseq (Graham et al. 2015). In these cases, optimising the preservation and extraction of DNA will need to be considered prior to sampling and DNA extraction. Choosing appropriate genetic markers and the method of isolation for studies of population genetic structure is also a non-trivial task. However, a number of comprehensive review papers have been published to address this question (Hellberg & Burton 2002; Schlötterer 2004; Allendorf et al. 2010; Cuéllar-Pinzón et al. 2016; Allendorf 2017). In addition, tools and papers exist that can help practitioners choose the appropriate number of samples and genetic markers for their particular question (e.g. Hoban et al. 2013). Prior to commencing development work, the literature should be screened thoroughly to determine whether genetic markers of a suitable resolution are already available for a candidate species – this can avoid the costs and time typically required for the development of novel markers. For example, SNP panels are now available for a wide range of marine species (e.g. salmonids, Meek et al. 2016; crustaceans, Jenkins et al. 2018b; and molluscs, Jiao et al. 2014), and are likely to be useful for the analysis of genetic structure, population assignment and connectivity.

3.3 Discussion

As this thesis focuses on assessing connectivity between British MPAs, species selected for further study had to occupy habitats that are located within the boundaries of MPAs designated in UK waters. Candidate species whose distribution also spans adjacent seas were desirable because this enabled spatial genetic structure and connectivity to be explored at a much broader scale (e.g. across neighbouring European seas). In effect, this meant that selected species had to have a northeast Atlantic distribution and be relatively abundant across part or most of the British Isles.

After consideration of the biological and methodological factors outlined previously (Table 4), two species were identified as suitable candidates: the pink sea fan (*Eunicella verrucosa*) and the European lobster (*Homarus gammarus*) (Fig. 13). However, it is important to be aware that the factors outlined in this framework differ in their ease of being assessed. For example, some factors may be prone to

subjectivity (e.g. ecological importance) and are potentially more difficult to assess, while other factors are somewhat objective (e.g. the availability of genetic markers) and are thus easier to assess. In addition, some factors in this framework can be more important than other factors for assessing genetic connectivity between MPAs. For instance, as mentioned above, to assess connectivity between a network of MPAs, the distribution of the candidate species must overlap that of MPA boundaries. When considering the relative importance of each factor on a study-by-study basis, a weighting criteria could be applied, such that each factor is ranked according to its importance for the study; this would facilitate an objective process of selecting (or omitting) taxa.

In this thesis, selecting the pink sea fan and the European lobster meant that population structure and genetic connectivity could be explored in two very different taxa with different life histories. The pink sea fan is a sessile gorgonian soft coral, is relatively long-lived with the potential for sexual and asexual reproduction, and is primarily found in habitats where there are hard substrates (i.e. rocky substrata) to attach to (section 4.1.1). In contrast, the European lobster is a sexually reproducing decapod crustacean with some capacity for movement as adults, and is found in habitats composed of hard and soft substrates where there is adequate shelter (section 6.1.1). Both species are broadcast spawners, but the development of eggs and larvae are markedly different, for example, pelagic larvae of pink sea fan are lecithotrophic whereas pelagic larvae of European lobster are planktotrophic. Choosing these two species enabled a multi-species assessment of connectivity between British MPAs to be conducted, with the potential to inform both marine conservation and fisheries management.

Based on the framework from Table 4, there were several advantages for selecting the pink sea fan (Table 5). Firstly, its distribution and abundance are sufficient to explore connectivity patterns within and outside of MPA boundaries. Secondly, it is generally considered a flagship species for UK conservation and is accordingly protected in England and Wales; it is also listed as 'Vulnerable' on the IUCN Red List. Thirdly, and perhaps most importantly for this candidate species, sampling in sufficient quantities had already taken place prior to this PhD. This meant there was no need to arrange permits for tissue collection (required in the UK because of the protective status of *E. verrucosa*), or organise sampling trips which would likely involve SCUBA diving and be very costly and time consuming.

Similarly, for the European lobster (*Homarus gammarus*), there were a number of advantages for selecting this species (Table 5). For example, although samples needed to be collected *de novo*, an active fishery across the British Isles and neighbouring seas meant there was massive potential to collect a vast number of

Figure 13: Species selected: the pink sea fan (*Eunicella verrucosa*) (left) and the European lobster (*Homarus gammarus*) (right).

samples from a variety of locations across its range. The commercial interest of this species also meant that the biology of *H. gammarus* is well known, which would likely help to explain some of the patterns observed from a novel genomics study. Moreover, *H. gammarus* is reasonably abundant across the British Isles and most of the northeast Atlantic, meaning connectivity patterns could be explored within and outside of MPA boundaries, but also across the northeast Atlantic and parts of the Mediterranean Sea.

In conclusion, this study has created a framework which highlights a number of biological and methodological factors to consider before selecting taxa to use in assessments of genetic connectivity between MPAs. After consideration of these factors, two species with differing life histories were selected to use as surrogates for assessing genetic connectivity between MPAs across the UK. In addition, because of the ecological and economical importance of these two species, the findings of these studies may also inform conservation or fisheries management bodies relevant to each species. Subsequently, the upcoming chapters explore the spatial genetic structure and connectivity of these two species using microsatellite and/or SNP markers.

Chapter 4: Population genetic structure and connectivity of pink sea fans (*Eunicella verrucosa*) using microsatellite and SNP markers

---

Microsatellite study

The microsatellite-based study of two octocoral species discussed in this chapter is based on a paper published in the journal *Heredity*. Tom L. Jenkins analysed the data, generated the figures and tables, wrote the first draft of the manuscript, and addressed all reviewer comments and edits. This study also formed part of the PhD thesis of Lyndsey P. Holland and, therefore, the results of this study are discussed as part of the Introduction and the Discussion as per university guidelines. The reference is given below and the full paper is available in the Appendix.

Holland LP*, Jenkins TL*, Stevens JR (2017) Contrasting patterns of population structure and gene flow facilitate exploration of connectivity in two widely distributed temperate octocorals. *Heredity* **119**, 35-48. *Joint first authorship.

SNP study

The preliminary SNP study was designed and carried out by Tom L. Jenkins and Jamie R. Stevens. SNPsaurus (Oregan) prepared the nextRAD libraries and identified SNPs using their custom bioinformatics pipeline.

4.0 Abstract

The pink sea fan (*Eunicella verrucosa*) is a priority species for conservation in English and Welsh waters; yet, until recently very little was known about its genetic diversity, population structure, dispersal and connectivity. In this chapter, the first population genetics study of *E. verrucosa* using 13 microsatellite markers is briefly discussed; this research spanned two PhD programmes (Lyndsey P. Holland and Tom L. Jenkins). Following this study, a novel preliminary study was carried out that took advantage of a relatively new reduced-representation sequencing (RRS) method, nextRAD, to isolate genome-wide single nucleotide polymorphisms (SNPs). The main aims of this novel study were to (i) test whether a RRS approach was feasible using the somewhat degraded DNA obtained from our *E. verrucosa* samples, and (ii) test if patterns of genetic diversity and population structure were comparable (or different) between microsatellite and SNP markers. The results suggested that isolating SNPs from across the *E. verrucosa* genome was feasible using nextRAD, although a handful of samples and one population had to be discarded due to the influence of missing data affecting downstream analyses. Spatial genetic patterns using 3,743 SNPs supported the results from the microsatellite study; three main genetic clusters were identified and organised into samples from Britain-France, western Ireland and southern Portugal, with evidence of weak differentiation between samples from Britain and France. Genetic diversity measures were relatively low and consistent across populations (agreeing with the microsatellite study); moreover, significant heterozygote deficiencies in all sampling sites were observed, possibly caused by inbreeding or a Wahlund effect. Overall, while it appears that microsatellite and SNP markers show similar patterns of genetic diversity and population structure in *E. verrucosa*, the inclusion of more individuals and intermediate sample sites for the SNP markers is necessary to fully validate these findings and to further explore the drivers of these patterns.

4.1 Introduction

4.1.1 Pink sea fan biology

The pink sea fan (*Eunicella verrucosa*) is a colonial gorgonian belonging to the class Anthozoa and the subclass Octocorallia. They are generally found on rocky substrates at depths of 10-150 m in areas of moderate to high water currents and are native across the northeast Atlantic and parts of the Mediterranean Sea where there is suitable habitat (Hayward & Ryland 1995). In particular, colonies have been recorded in western Africa and the western Mediterranean (southern range),

around the coasts of Portugal, northern Spain and northwest France, and up to southwest Britain and northwest Ireland (northern range). Fully-grown colonies usually stand about 30 cm tall but they can reach up to 75 cm in some areas (Wood 2013). Colonies are commonly seen orientated towards the direction of ocean currents to allow the polyps to filter nutrients or prey using their tentacles. Two distinct colour morphs (orange-pink and white) of *E. verrucosa* have been found (Fig. 14), though it is not known whether one phenotype carries any fitness advantage for colonies.

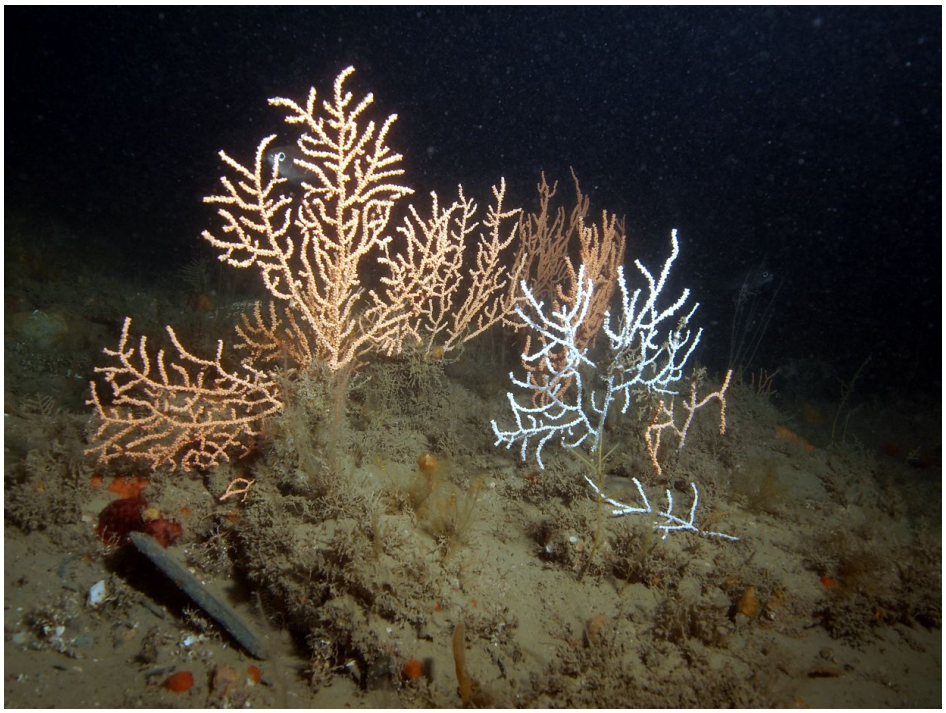


Figure 14: Pink sea fan (*Eunicella verrucosa*) orange-pink (left) and white (right) colour morphs (image provided by Chris Wood).

Eunicella verrucosa is thought to be a gonochoristic (separate sexes) and sexually reproducing species, yet asexual reproduction may be possible by clonal fragmentation (Munro 2004). Colonies are broadcast spawners, releasing gametes into the water column towards the end of summer (August-September), which are externally fertilised (Munro 2004). Larvae are thought to be lecithotrophic, meaning they are provided with a yolk sac as a source of nutrition to use during their dispersal via ocean currents. However, the pelagic larval duration (PLD) of *E. verrucosa* is unknown, which presents uncertainty over the dispersal capability of this species.

In many sublittorial ecosystems, *E. verrucosa* colonies can indirectly or directly support other marine organisms, particularly when they are locally abundant and form large 'forests' (Wood 2013; Pikesley et al. 2016). These forests (Fig. 15)

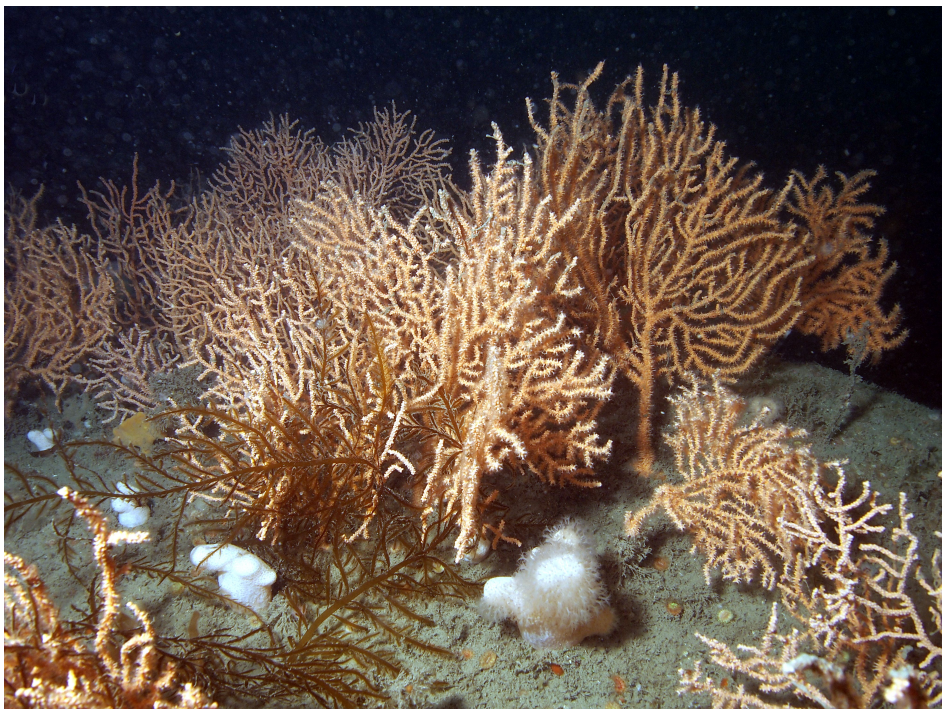


Figure 15: A pink sea fan (*Eunicella verrucosa*) 'forest' (image provided by Chris Wood).

provide structural complexity and habitat for a number of epifaunal animals and likely play an important sheltering role for small or juvenile organisms seeking to take refuge from predators. In addition, some organisms are known to settle on (e.g. barnacles and bryozoans), attach eggs on (e.g. catsharks), or even exclusively live on *E. verrucosa* and other sea fans (e.g. the sea fan anemone, *Amphianthus dohrnii*) (Wood 2013). This suggests that *E. verrucosa* plays an important role in the functional ecology of the benthic communities it resides in and could be considered an ecosystem engineer (Hall-Spencer et al. 2007; Pikesley et al. 2016).

4.1.2 Conservation status

Pink sea fans are extremely vulnerable to seabed disturbance from trawling and other gears and marine litter (Hinz et al. 2011; Sheehan et al. 2017), primarily because of their delicate structure and overall slow growth rates ($\sim$3.33 cm year^{-1} when colony height $<$15 cm and $\sim$0.62 cm year^{-1} when colony height $>$40 cm, Sartoretto & Francour 2012). Accordingly, the pink sea fan has been classified as 'Vulnerable' by the IUCN Red List since 1996, which is defined as a species facing a very high risk of extinction in the wild in the medium-term future. In addition, *E. verrucosa* is also listed as a priority species under the UK Biodiversity Action Plan and a species of principal importance in England under the NERC Act 2006. In response, the UK government has established several MCZs around southwest

England (e.g. Chesil Beach and Stennis Ledges, Isles of Scilly, Skerries Bank and Surrounds, The Manacles, and Whitsand and Looe Bay) and around Wales (e.g. Skomer Island), that specifically identify *E. verrucosa* as a protected feature in their designation listing. To my knowledge, pink sea fans are not explicitly protected outside of the UK. Nevertheless, the Republic of Ireland have established SACs under the EU Habitats Directive (e.g. Kenmare River SAC, Galway Bay Complex SAC and Donegal Bay SAC) and France has designated OSPAR MPAs in northwest Brittany (e.g. the Glenan Islands and the Baie de Morlaix, OSPAR Commission 2016) where *E. verrucosa* colonies are known to occur, which may indirectly help to protect colonies from disturbance in those areas.

4.1.3 Previous genetic research

Genetic research on pink sea fans has been scarce, particularly at the population level. Studies involving *E. verrucosa* have tended to focus on comparative transcriptomics across diverse taxa (Romiguier et al. 2014) or on phylogenetic relationships within the Anthozoa (Pratlong et al. 2016) or within the *Eunicella* genus (Aurelle et al. 2017). As a result, not much is known about the genetic diversity and population structure of this species, the study of which may shed light on patterns of connectivity among populations and the dispersal capacity of larvae.

The first population genetic study of *E. verrucosa* was published in early 2017 (Holland et al. 2017). In this study, the genetic diversity and population structure of *E. verrucosa* and another octocoral, dead man's fingers (*Alcyonium digitatum*), were explored across the northeast Atlantic using 13 (*E. verrucosa*) and eight (*A. digitatum*) microsatellite loci developed by Holland et al. (2013a, 2013b). Details of DNA extraction, microsatellite development and scoring, and analytical methods can be found in the original papers (Holland et al. 2013a, 2013b, 2017). Due to the focus of this chapter on *E. verrucosa*, hereafter only the results for *E. verrucosa* are discussed.

Tissue samples of *E. verrucosa* were collected in 2007-2012 from sites across its middle (southern Portugal) and northern range (southwest Britain/northwest Ireland) (Fig. 16). This study revealed several important findings about the population genetics of *E. verrucosa*. Firstly, genetic diversity (expected heterozygosity and allelic richness) was generally uniform across the sampling range, but was lower compared to that reported in other temperate corals (Table 3 in Holland et al. 2017), which included the closely related species *E. cavolini* (Masmoudi et al. 2016) and *E. singularis* (Costantini et al. 2016), *A. digitatum* (Holland et al. 2017), and two Mediterranean octocorals, *Corallium rubrum* (Ledoux

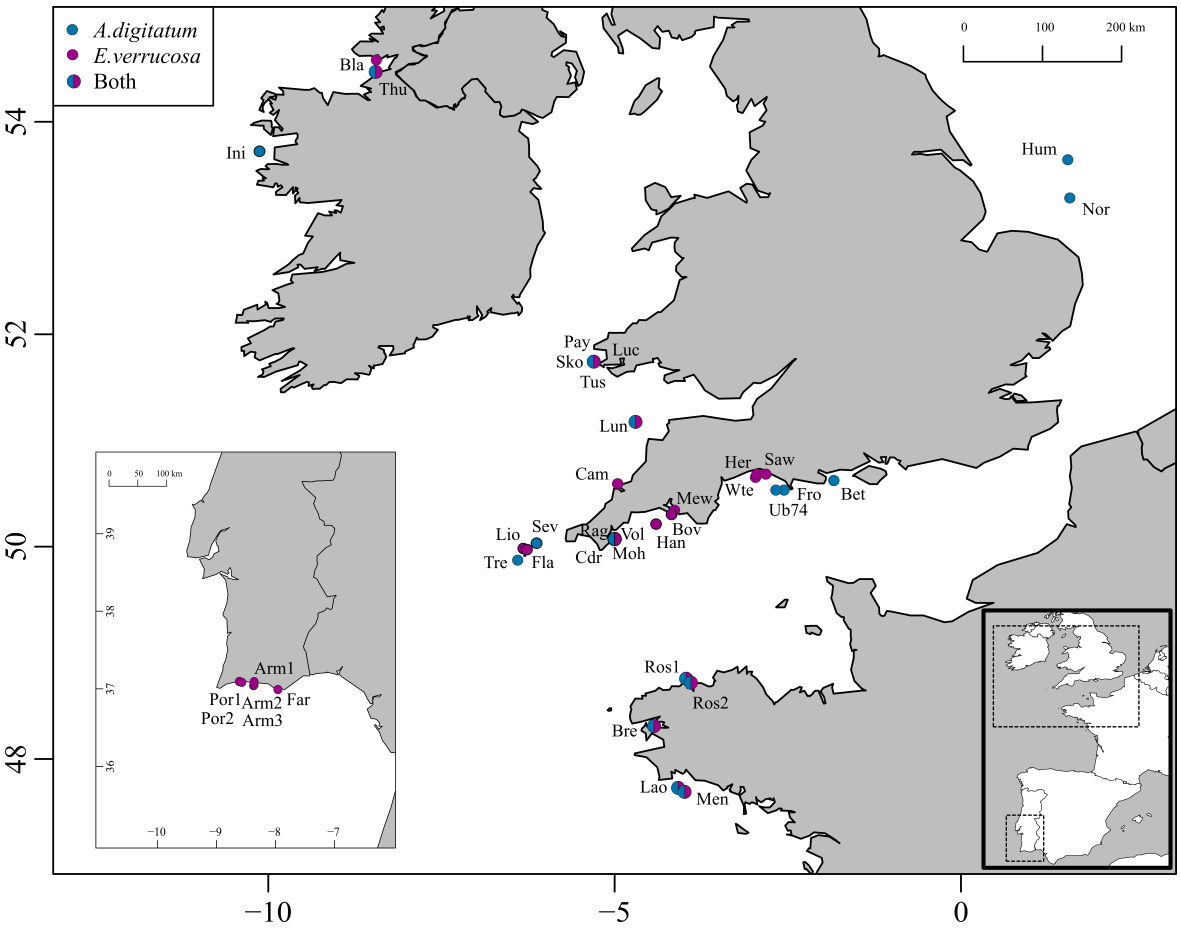


Figure 16: Map of the sampling sites used in the microsatellite study from Holland et al. (2017).

et al. 2010) and *Paramuricea clavata* (Mokhtar-Jamai et al. 2011). However, in comparison to these temperate octocorals, the number of sites with significant heterozygotes deficiencies was lower in both *E. verrucosa* and *A. digitatum* (Holland et al. 2017). The authors concluded that the most likely explanation for this pattern was that, overall, the frequency of inbreeding is generally low in *E. verrucosa* and *A. digitatum*, though inbreeding may be apparent at some sites which is likely due to site-specific factors (Holland et al. 2017).

Secondly, across the sites sampled, *E. verrucosa* populations were not panmictic, instead revealing three distinct genetic clusters (Britain-France, southern Portugal and northwest Ireland) (Fig. 17), with evidence of weak genetic differentiation between sites from southwest Britain and northwest France (Fig. 17). Further analysis suggested that isolation-by-distance (IBD) was a likely explanation for the differentiation observed between sites from Britain, France and Portugal. Interestingly, this pattern of IBD appears to be common in temperate octocorals (Table 3 in Holland et al. 2017), which is possibly due to their sedentary life history and their lack of, or shorter, PLD compared to other benthic marine species.

However, IBD did not explain the genetically distinct profiles observed in colonies from northwest Ireland. This suggested that the differentiation observed in northwest Ireland colonies, inhabiting the northern peripheral range of the species, may be driven by other factors, such as barriers to gene flow and/or selection (Holland et al. 2017). Two previous studies of marine invertebrates studied across this region have also reported genetic differentiation in western Ireland compared to other locations in the northeast Atlantic (Remerie et al. 2009; Casu et al. 2011). These studies attributed this differentiation to recolonisation from glacial refugia or from persistance in ice-free coastal areas during the LGM. Lower genetic diversity at range margins can be explained by founder effects and genetic drift (due to a low number of post-glacial recolonisers); Casu et al. (2011) concluded this was the likely explanation in their study of microturbellarians from southwest Ireland. In contrast, Remerie et al. (2009) found higher genetic diversity and heterogeneity for *Neomysis integer* (an estuarine shrimp) in glaciated areas, suggestive of range persistence during the LGM. The findings for *E. verrucosa* colonies from northwest Ireland, which exhibited the lowest genetic diversity detected in the entire study *E. verrucosa*, are in line with those of Casu et al. (2011); this suggests that founder effects following post-glacial recolonisations are a potential explanation for the genetic distinctiveness of northeast Ireland *E. verrucosa* colonies. Yet, insufficient sampling across the rest of Ireland and at the southern-most limits of the range of *E. verrucosa* makes inferences about the origin of populations in northwest Ireland difficult. In addition, lower diversity at the range margins can be characteristic of populations under intense selection pressures (Johannesson & André 2006), which may suggest that natural selection may be driving this genetic divergence; however, it is not known which selection pressures, if any, may be acting on these most northerly populations of pink sea fan.

Thirdly, there was strong genetic similarity within regions (i.e. within southwest Britain), suggestive of high genetic connectivity at these spatial scales (Fig. 17). This was further supported by analyses of contemporary gene flow, which indicated that the majority of gene flow was exchanged between sites from the same region (Fig. 18). In addition, this analysis also provided evidence that colonies from southwest Britain have potentially been a source of genetic variants for French colonies over the last few generations. Together, these results suggest that *E. verrucosa* larvae are able to disperse and exchange genetic material at distances of up to 500 km; however, whether gene flow at these scales are be achieved by a single migration event or by a stepping-stone model of connectivity is not yet clear. Of course, this genetic similarity may also suggest that at a regional scale *E. verrucosa* have high effective population sizes, which can hinder the

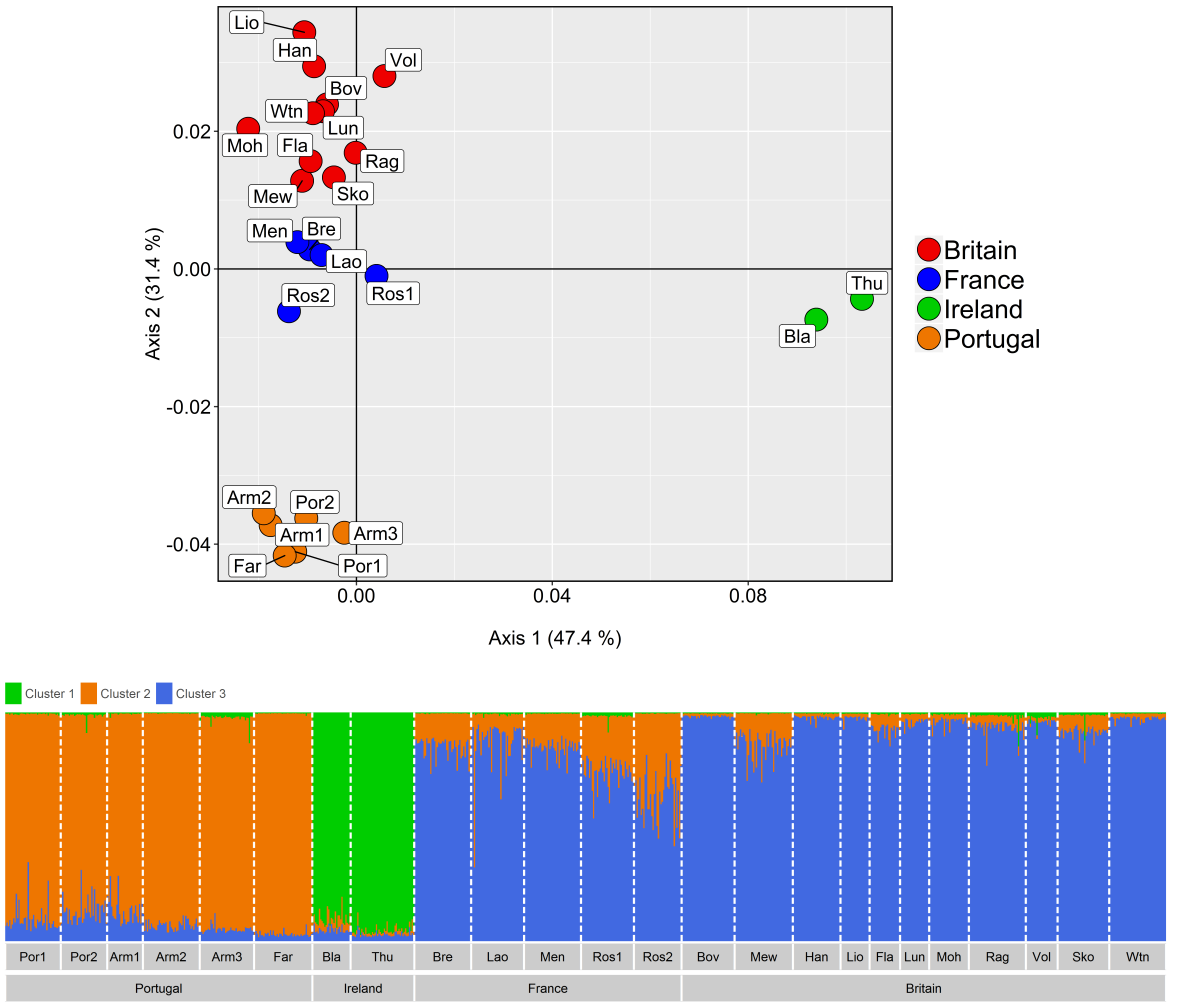


Figure 17: Population structure of pink sea fans using 13 microsatellite markers (Holland et al. 2017). A principle coordinates analysis (top) and results from STRUCTURE analysis (bottom) are presented. For the principle coordinates analysis, each point represents a sampling site and the colours correspond to the country of origin. For the STRUCTURE plot, each bar represents an individual and the colours represent the membership proportion to each of the three genetic clusters inferred.

formation of population structure by mitigating the influence of drift.

4.1.4 Study aims

The microsatellite study summarised previously provided a much-needed investigation into the population genetics and connectivity of pink sea fans across their middle and northern range. However, as a Red Listed and priority species for conservation, it is critical that genetic data presented to MPA managers are as robust and reliable as possible before decisions are made and enacted. Therefore, a study that explores the genetic patterns of *E. verrucosa* using an alternative marker system is needed, which would facilitate a direct comparison with the 13

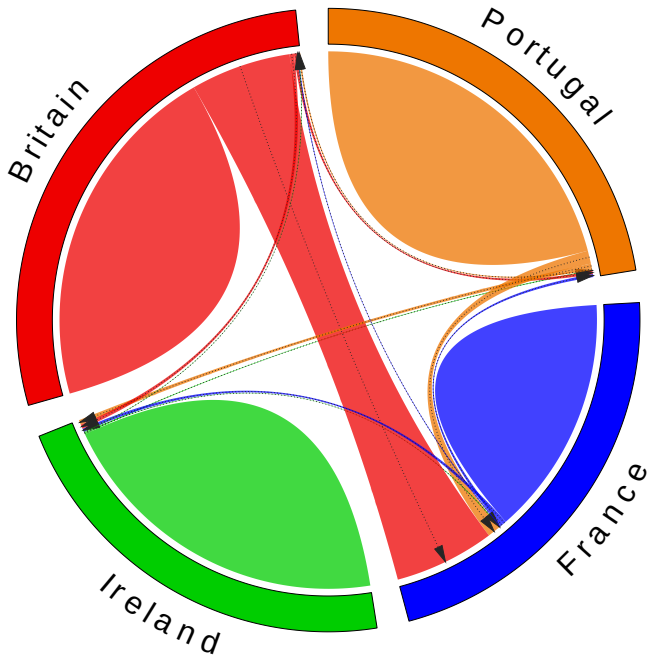


Figure 18: *Eunicella verrucosa* gene flow using 13 microsatellite markers (Holland et al. 2017). Gene flow was calculated using BayesAss v3.0.4 (Wilson & Rannala 2003), which provides an estimate of gene flow over the last few generations. For this analysis, sampling sites in each country were combined (denoted by colours). The direction of an arrow represents the direction of gene flow from one country to another; the width of the arrows denotes the relative amount of gene flow (i.e. the wider the arrow, the more gene flow). The 'humps' represent gene flow originating from sample sites within countries.

microsatellite markers.

In light of this, the first aim was to conduct a preliminary study to investigate whether isolating genome-wide SNPs using a reduced representation sequencing (RRS) approach was feasible. As the quality of DNA required for microsatellite genotyping does not necessarily need to be of high molecular weight and quality, it was essential to re-extract and optimise a DNA extraction protocol for *E. verrucosa* samples. This was because the DNA samples used in Holland et al. (2017) were relatively degraded, proving inadequate for high-throughput sequencing (K. Moore, Exeter Sequencing Service, *pers. comm.*). However, even after trying many different extraction methods, only suboptimal degraded DNA could be extracted for some *E. verrucosa* samples. The second aim of this study was to use these preliminary data to test whether thousands of SNP markers showed similar patterns of spatial genetic structure to the 13 microsatellite markers analysed in Holland et al. (2017). Finally, to conclude, future research objectives and directions for the population genomics study of *E. verrucosa* are discussed.

4.2 Materials and methods

4.2.1 Sample collection and DNA extraction

Eunicella verrucosa tissue samples were collected and preserved in 95-100 % ethanol as described in Holland et al. (2017). Eight sampling sites from Holland et al. (2017) were chosen to include in this study and an additional novel site from the Isles of Scilly (Hathor Wreck) was also included (Table 6; Fig. 19). All samples used for this study were collected between 2009 and 2012. As one of the aims of this study was to compare the resolution obtainable with SNPs to microsatellite markers, the sampling sites and number of individuals were chosen such that roughly equal representation from the three genetic groups found in Holland et al. (2017) were included.

Genomic DNA was extracted from 15-20 polyps using a modified salting-out protocol (Appendix A4), originally designed to extract DNA from crustacean exoskeletal tissue (Li et al. 2011). Obtaining high molecular weight and pure DNA from many of the pink sea fan tissue samples was difficult, with evidence of degradation present in nearly every sample (Appendix A5). Initially, the Blood and Tissue Kit (Qiagen) was used for extractions; however, these DNA samples failed to generate adequate RAD libraries during a trial run and comparisons suggested this method was inferior to the salting-out protocol (Appendix A5). The concentration and purity of all extractions were quantified by spectrophotometry using a NanoDrop 1000. In addition, all DNA samples were further evaluated by running the DNA on a 1 % agarose gel and by quantifying their concentration with fluorometry using the Invitrogen Qubit Assay kit, which only measures the amount of double-stranded DNA in the sample. These quality assessments were vital to allow the best quality DNA samples from each sampling site to be submitted for high-throughput sequencing.

4.2.2 nextRAD sequencing

A relatively new RRS method, nextRAD (Russello et al. 2015), was chosen for isolating SNPs from across the *E. verrucosa* genome. The rationale behind this choice was that only 10 ng or less is required as input DNA and the method was suggested to perform better than traditional RADseq with degraded DNA (SNPsaurus, *pers. comm.*). Genomic DNA was converted into nextRAD genotyping-by-sequencing libraries (SNPsaurus, LLC) which uses a selective primer sequence (rather than restriction enzymes) to genotype loci consistently across samples (Russello et al. 2015). Genomic DNA was first fragmented with

Table 6: Pink sea fan (*Eunicella verrucosa*) sampling information for the SNP study.

Country	Site	Code	*N*	Lat	Lon	Depth (m)	Year
Britain	[a]Isles of Scilly, Hathor Wreck	Hat	9	49.88	-6.35	28	2010
	[a]Isles of Scilly, Lion Rock	Lio	9	49.98	-6.31	24	2009
	[a]Lundy Island	Lun	18	51.17	-4.69	23	2009
France	[b]Glenan Islands, Laonegued	Lao	10	47.73	-4.06	30	2011
	[b]Glenan Islands, Men Goe	Men	9	47.69	-3.99	30	2011
Ireland	[b]Donegal Bay, Black Rock	Bla	10	54.58	-8.43	25	2012
	[b]Donegal Bay, Thumb Rock	Thu	10	54.47	-8.44	20	2012
Portugal	Faro	Far	10	36.98	-7.99	17	2010
	Portimao2	Por	10	37.10	-8.56	18	2010

N, number of individuals submitted for sequencing.
[a]Marine Conservation Zone.
[b]Special Area of Conservation.

Nextera reagent (Illumina, Inc.), which also ligates short adapter sequences to the ends of the fragments. The Nextera reaction was scaled for fragmenting 5 ng of genomic DNA, although 7.5 ng of genomic DNA was used for input to compensate for the amount of degraded DNA in the samples and to increase fragment sizes. Fragmented DNA was then amplified for 26 cycles at 73°C, with one of the primers matching the adapter and extending nine nucleotides into the genomic DNA with the selective sequence GTGTAGAGG. Thus, only fragments starting with a sequence that can be hybridized by the selective sequence of the primer will be efficiently amplified. The nextRAD libraries were sequenced on one Illumina HiSeq 4000 lane with single-end 150 bp reads (University of Oregon). The quality of reads for each sample was assessed using FastQC software (Babraham Bioinformatics).

4.2.3 Bioinformatics

The genotyping analysis used custom scripts (SNPsaurus, LLC) that trimmed the reads using bbduk (BBMap tools). A *de novo* reference was created by collecting 10 million reads in total, evenly from the samples, and excluding reads that had counts fewer than seven or more than 700. The remaining loci were then aligned to each other to identify allelic loci and collapse allelic haplotypes to a single representative. All reads were mapped to the reference with an alignment identity threshold of 93 % using bbmap (BBMap tools). Genotype calling was done using Samtools and bcftools. The output vcf file was filtered to remove alleles with a population frequency of less than 3 %. Loci were removed that were heterozygous in all samples or had more than two alleles in a sample (suggesting collapsed paralogs). The absence of artefacts was checked by counting SNPs at each nucleotide position and ensuring that the number of SNPs did not increase with reduced base quality at the end of

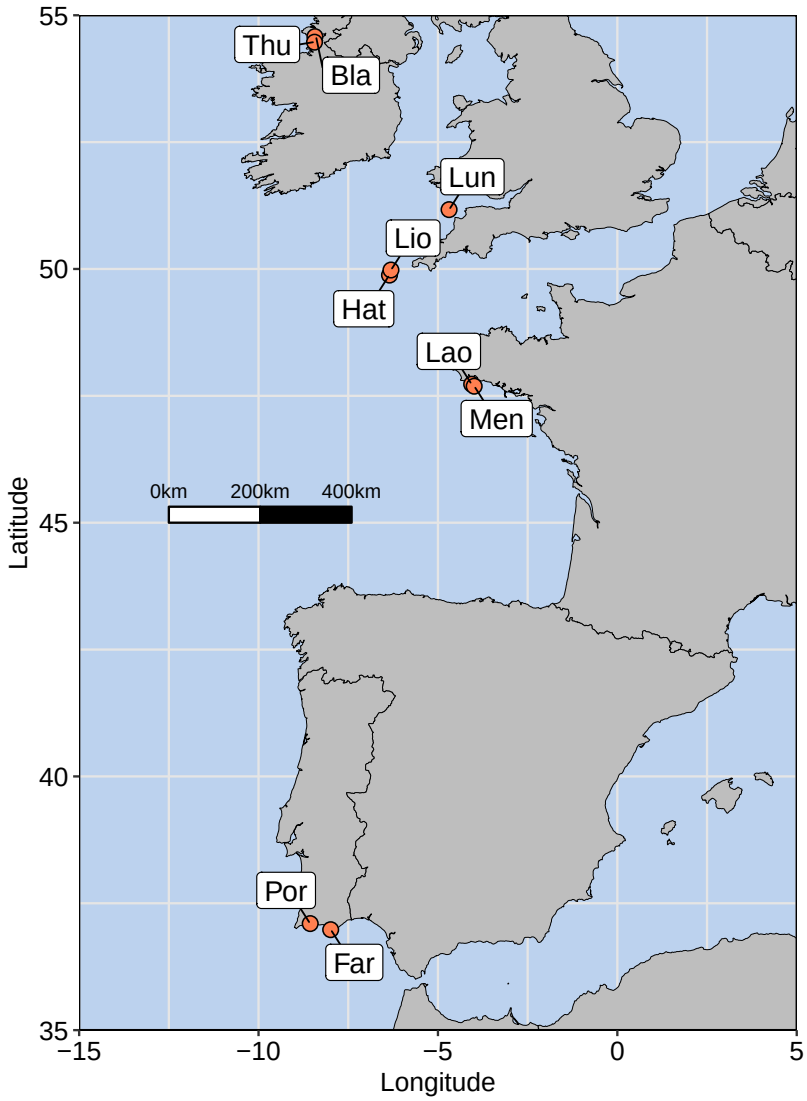


Figure 19: *Eunicella verrucosa* sampling sites for the SNP study. See Table 6 for sampling site code information.

the read.

4.2.4 Quality control and filtering

Filtering of the vcf file provided by SNPsaurus was implemented in Stacks v1.48 (Catchen et al. 2013) and radiator v0.0.5 (Gosselin 2017). The `populations` program in Stacks was run with the following parameters: a locus had to be present in at least 80 % of individuals in a population (`-r` 0.8) and in at least seven (out of nine) populations (`-M` 7), the minor allele frequency threshold was set to 5 % (`-min_maf 0.05`), loci had to have a maximum observed heterozygosity of less than 0.5 (`-max_obs_het 0.5`), and only the first SNP in each locus was processed (`-write_single_snp`). A new vcf file was created and then imported into R for

further filtering using radiator; several functions were run to manipulate and interrogate the data. Firstly, the `missing_visualization` function was executed using default parameters to visualise and assess missing data. Secondly, the `detect_mixed_genomes` function was run using default parameters which assesses observed heterozygosity in each individual for a diagnostic test of mixed samples or poor polymorphism discovery. Thirdly, the `detect_duplicate_genomes` function was run with the genome argument set to true (`genome=TRUE`) to highlight potential duplicate individuals. After data exploration, only loci that were genotyped in all populations (`common.markers=TRUE`) were retained. The final filtered dataset was exported in multiple formats using the `genomic_converter` function.

4.2.5 Genetic diversity

Observed heterozygosity (H_o), expected heterozygosity (H_e) and the inbreeding coefficient (F_{is}) were calculated using the `divBasic` function from the R package diveRsity v1.9.90 (Keenan et al. 2013). For each population, the mean value across all loci was computed and significance of F_{is} was assessed by calculating bias corrected 95 % confidence intervals (1,000 bootstrap replicates) and testing whether values were significantly different from zero.

4.2.6 Population structure

Genetic differentiation between sampling sites was analysed by calculating pairwise values of F_{st} (Weir & Cockerham 1984) and D (Jost 2008) using the `diffCalc` function from diveRsity. Heatmaps of each statistic were visualised in R and significance was assessed using the same approach previously described for F_{is}.

Population structure among sampling sites was explored using two different approaches. Firstly, a discriminant analysis of principal components (DAPC) was run using the `dapc` function from the R package adegenet v2.1.0 (Jombart & Ahmed 2011). DAPC attempts to summarise genetic differentiation between groups, sampling sites in this context, while overlooking variation within groups (Jombart et al. 2010). It first transforms the data using principal components analysis (PCA) to ensure variables are uncorrelated and that the number of variables (alleles) is less than the number of observations (individuals) in the dataset. These are necessary prerequisites for discriminant analysis (DA). Then a DA is performed on the number of principal components (PCs) retained, which is selected by the user. Cross validation using the `xvalDapc` function from adegenet was used to choose the optimal number of PCs to retain.

Secondly, STRUCTURE v2.3.4 (Pritchard et al. 2000), a Bayesian clustering algorithm, was run in parallel using the program StrAuto v1.0 (Chhatre & Emerson 2017). STRUCTURE attempts to estimate the number of ancestral populations (*K*) from multi-locus allele frequencies. Given a value of *K*, the program assigns individuals probabilistically to these *K* clusters, with the assumption that loci are under Hardy-Weinberg equilibrium (HWE) and linkage equilibrium (Gilbert 2016). Post-hoc tests have been developed to help assess the *K* that best fits the data (e.g. mean L(*K*), Pritchard et al. 2000; and delta *K*, Evanno et al. 2005); however, several recent papers have suggested that detecting a true *K* statistically is extremely difficult and that in some cases no true *K* may exist (Meirmans 2015; Gilbert 2016; Janes et al. 2017). The authors encourage visualising each *K* to explore population structure at each level to make informed choices. STRUCTURE was executed using the admixture model, with 10^5 MCMC repetitions and a burn-in of 10^5. The locprior option was selected, which meant that sampling site origins were used as priors; all other parameters were set to default values. The maximum assumed (*K*) was nine and ten independent replicates per *K* (1-9) were computed. The mean L(*K*) and delta *K* statistics were examined using the R package pophelper v2.2.5.1 (Francis 2017). Replicates runs were aligned and merged with CLUMPP (Jakobsson & Rosenberg 2007) using a wrapper script in pophelper and R was used to visualise the results.

4.2.7 Detecting outlier SNPs

To detect outlier SNPs potentially under selection, four differentiation-based methods were implemented: Arlequin v3.5.2.2 (Excoffier & Lischer 2010), Bayescan v2.1 (Foll & Gaggiotti 2008), OutFLANK v0.2 (Whitlock & Lotterhos 2015) and PCadapt v4.0.3 (Luu et al. 2017). Arlequin integrates heterozygosity and simulates a distribution of F_{st} for neutrally distributed markers to detect outliers. The infinite island model was run using 100,000 simulations and 1,000 demes. Bayescan is a Bayesian method based on a logistic regression model that attempts to distinguish locus-specific (alpha) effects of selection from population-specific (beta) effects of demography; departure from neutrality at a given locus is assumed when the locus-specific component is required to explain the observed pattern of diversity (Foll & Gaggiotti 2008). Bayescan was run using default parameters and a prior odds of 10,000, which sets the neutral model as being 10,000 times more likely than the model of selection to try and minimise the risk of false positives (Whitlock & Lotterhos 2015). OutFLANK calculates a likelihood on a trimmed distribution of F_{st} values to infer the distribution of F_{st} for

neutral markers and was executed using default parameters. PCAdapt uses principle components analysis to detect loci under selection and assumes that markers excessively related to population structure are candidates for local adaptation; default parameters were selected and three principle components were retained based on the number of clusters inferred from the DAPC (see Results). For all methods, a false discovery rate of 0.05 was used to identify outliers and only SNPs that were identified as outliers in two or more methods were classed as outlier SNPs. The DAPC was re-run with putatively neutral SNPs and SNPs putatively under divergent selection to explore the contribution of neutral versus potential adaptive processes in driving the genetic patterns observed.

4.3 Results

4.3.1 Sequencing

In total, 362 million reads were generated by the Illumina HiSeq 4000. The Phred score for most base calls exceeded 32, indicating that the sequence quality for reads across all samples was very good ($>$ 99.9 % accuracy). The number of reads was vastly unequally distributed among individual samples but there was no evidence to suggest that this pattern was population-specific (Fig. 20).

4.3.2 Quality control and filtering

The vcf file provided by SNPsaurus contained 18,459 loci which was reduced to 6,386 loci after filtering in Stacks. Analysis in radiator highlighted some potential problems with missing data and heterozygosity of loci. Firstly, there was greater than 30 % missing data in 11 samples, including Bla09, Lun11 and all samples from Men Goe (Fig. 21). Initial genetic analysis using STRUCTURE distinguished only Men Goe from all other samples; this finding does not accord with any obvious underlying biological, geographical or physical patterns or parameters, so this population was immediately removed from the analysis due to the potential bias of missing data. The remaining individuals generally showed less than 20 % missing data; therefore, individuals that met this threshold were retained to ensure downstream analyses would not be biased by missing genotypes. Similarly, enforcing a 20 % threshold for missing data for each locus ensured that filtering of loci was also stringent and that only the highest quality SNPs were retained.

Secondly, observed heterozygosity averaged across all loci for each individual showed pronounced differences relative to the mean of the population (Fig. 22). This pattern can suggest that samples have potentially been mixed or that

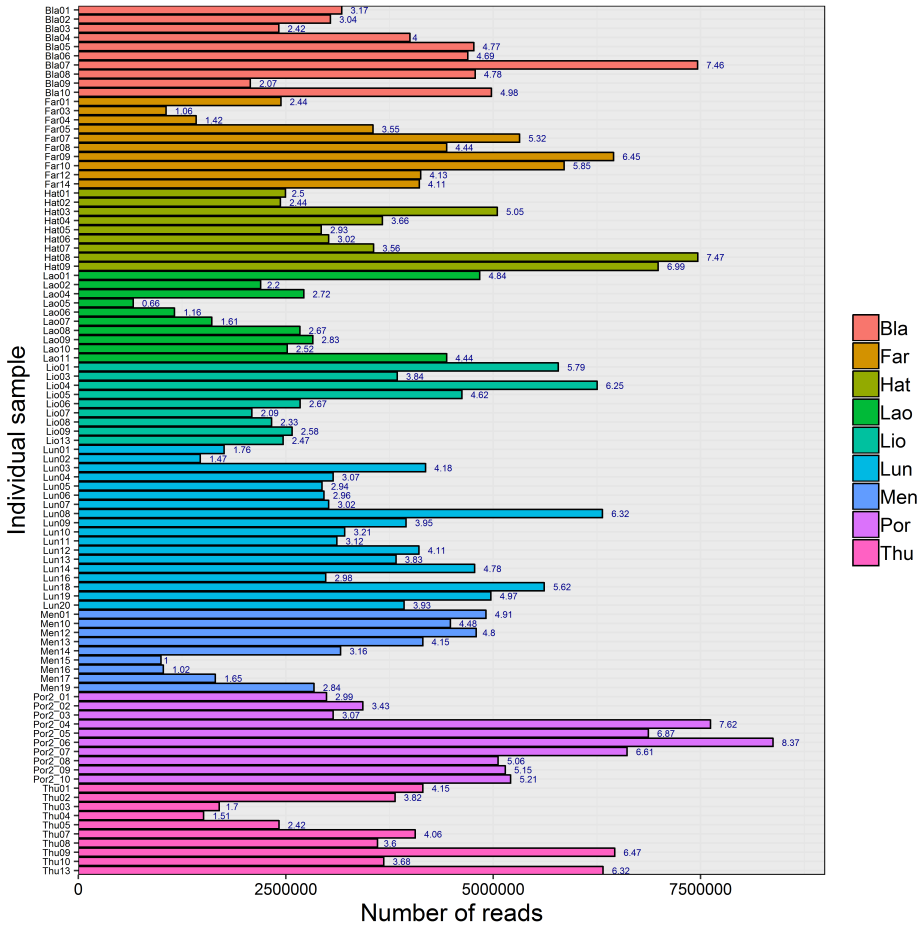


Figure 20: The number of reads per individual of *Eunicella verrucosa*. Colours denote the sampling site of origin for each individual.

polymorphism discovery has been biased based on DNA quality or other artefacts related to library preparation and sequencing (Gosselin 2017). Although it may be possible that bias could have been introduced into the dataset because of the degraded DNA of some of these samples, downstream genetic analyses made biological sense when using those loci retained post-filtering; therefore, no individuals or loci were omitted based on heterozygosity. Finally, comparison of two specimens, Lun03 and Lun10, revealed that they have more than 83 % of their genotypes in common (Fig. 23); therefore, only one sample was kept for further analyses because they could be duplicates. The final *E. verrucosa* dataset comprised 77 individuals, 8 sites and 3,743 SNPs.

4.3.3 Genetic diversity

Genetic diversity was relatively consistent across all populations. H_o and H_e ranged from 0.17-0.20 and 0.24-0.26, respectively (Table 7). Values of F_{is} for all populations were significantly positive (confidence intervals did not span zero), suggesting a

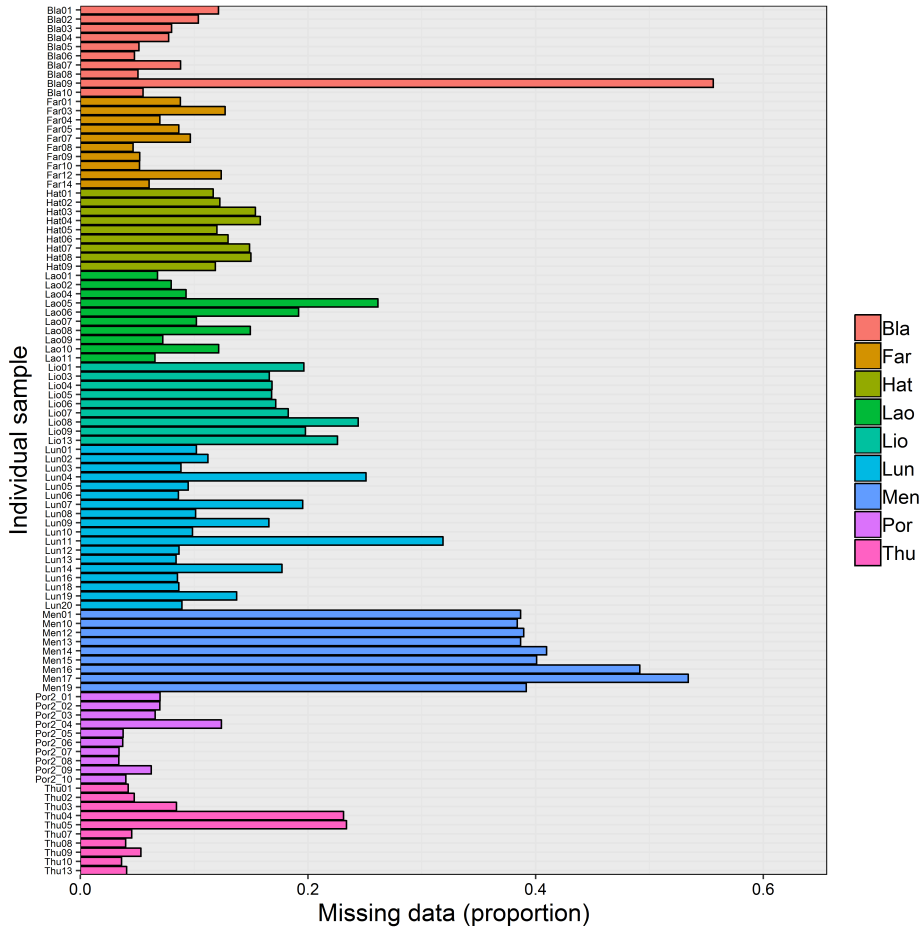


Figure 21: Proportion of missing data per individual of *Eunicella verrucosa*. Colours denote the sampling site of origin for each individual.

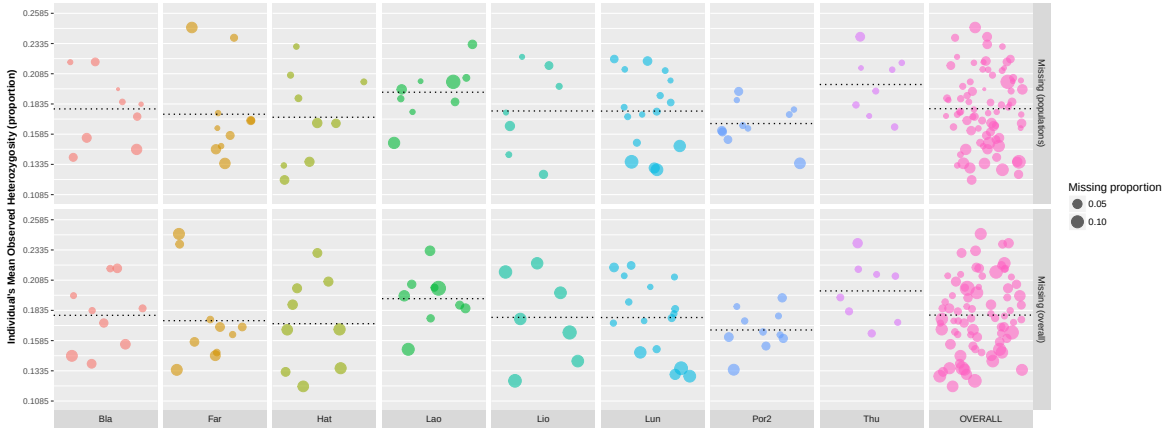


Figure 22: Observed heterozygosity per individual averaged across all loci. Each point represents an individual and the size of points is proportional to the amount of missing data for that individual. Colours denote sampling sites and the black dotted line represents the average heterozygosity for each sampling site.

profound deficiency in heterozygotes. Results were also comparable when only putatively neutral SNPs were used to calculate genetic diversity statistics (Table 7).

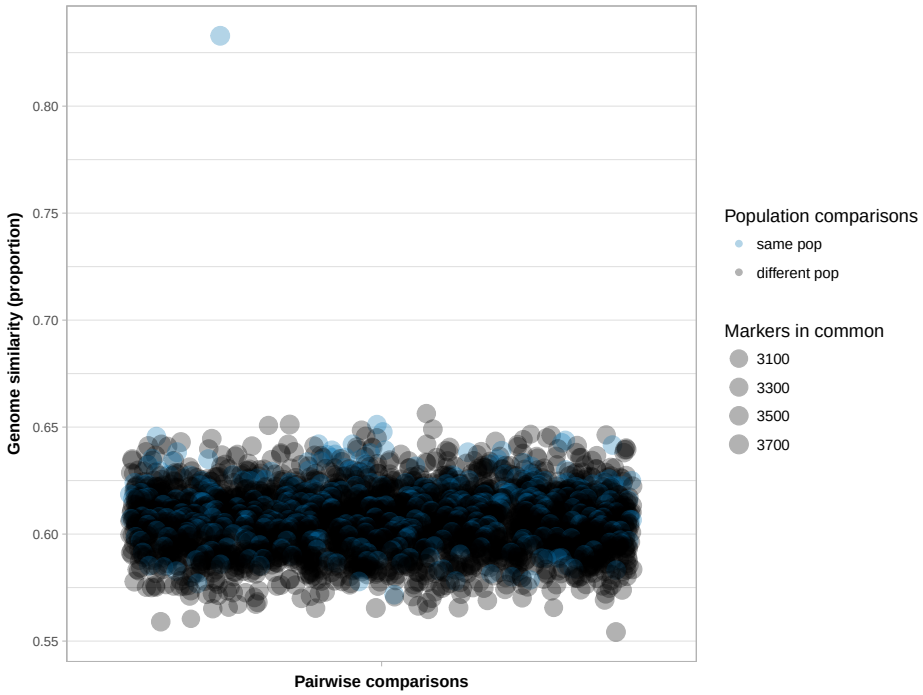


Figure 23: Individual pairwise genome similarity. The proportion of shared genotypes is averaged across shared markers between each pairwise comparison. Each point represents an individual and the size of points is proportional to the number of shared markers. Black points denote pairwise comparisons between different sampling sites and blue points denotes pairwise comparisons between individuals from the same sampling site.

Table 7: *Eunicella verrucosa* summary information and genetic diversity statistics.

Country	Code	N	N_g	H_o	H_e	F_{is}	$H_o^{neutral}$	$H_e^{neutral}$	$F_{is}^{neutral}$
Britain	Hat	9	9	0.17	0.25	**0.321**	0.17	0.28	**0.303**
	Lio	9	7	0.18	0.24	**0.271**	0.18	0.27	**0.267**
	Lun	18	15	0.18	0.26	**0.310**	0.18	0.27	**0.280**
France	Lao	10	9	0.19	0.25	**0.322**	0.19	0.27	**0.228**
Ireland	Bla	10	9	0.18	0.25	**0.280**	0.18	0.27	**0.264**
	Thu	10	8	0.20	0.25	**0.211**	0.20	0.27	**0.214**
Portugal	Far	10	10	0.17	0.25	**0.294**	0.17	0.27	**0.269**
	Por	10	10	0.17	0.25	**0.324**	0.17	0.27	**0.303**

N number of individuals submitted for sequencing, N_g number of individuals successfully genotyped, H_o observed heterozygosity, H_e expected heterozygosity, F_{is} inbreeding coefficient, ^{neutral} only putatively neutral SNPs were used for this calculation.
Bold font represents values that are significantly different from zero.

4.3.4 Population structure

Global F_{st} and D were 0.016 and 0.003, respectively, and both pairwise differentiation statistics showed comparable pairwise patterns between sampling sites (Fig. 24). Values of F_{st} ranged from zero (Bla-Thu) to 0.030 (Far-Bla) and from 0.006 (Bla-Thu) to 0.019 (Far-Bla) for D. The highest values for both statistics were between Portugal (Far and Por) and Ireland (Bla and Thu), whereas the

lowest values tended to be between sites originating from the same country which are spatially situated closer together. No values of F_{st} or D were significantly different from zero.

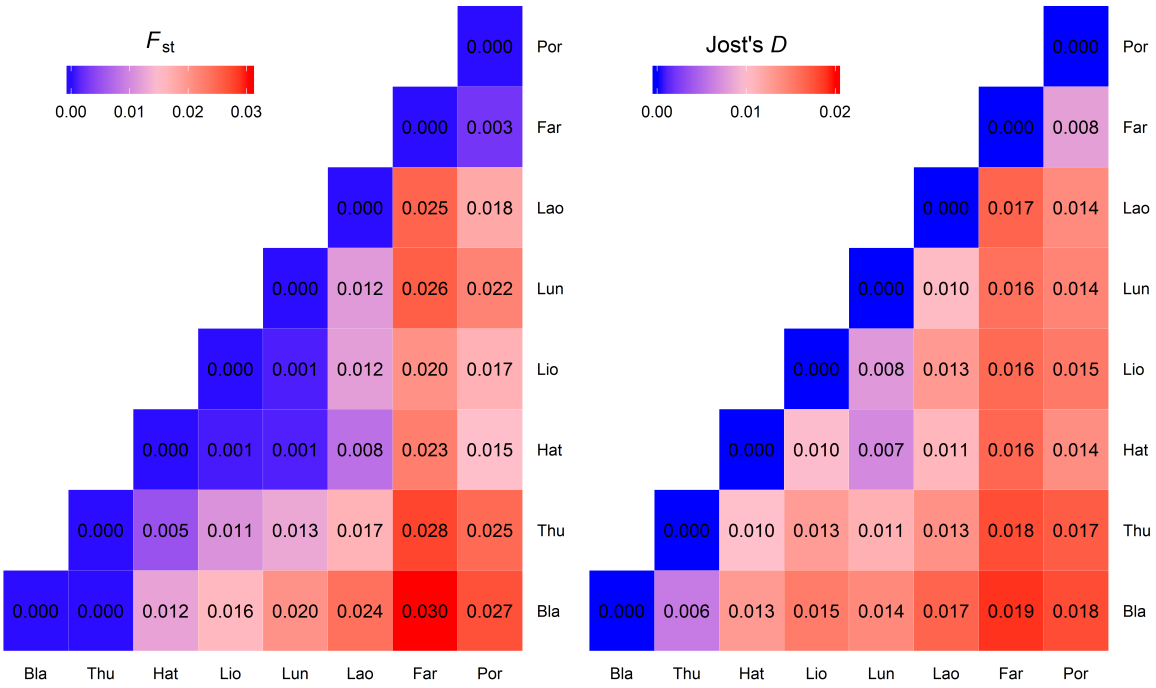


Figure 24: Pairwise genetic differentiation of Weir and Cockerham's F_{st} (left) and Jost's D (right) among sampling sites.

The DAPC using all SNPs provided strong support for three genetic groups: western Ireland (Bla and Thu), Britain-France (Hat, Lio, Lun and Lao) and southern Portugal (Far and Por) (Fig. 25). In total, axis 1 and axis 2 explained 88.7 % of the genetic variation. Laonegued appeared to marginally differentiate itself from the main Britain-France cluster, which suggests there is weak differentiation between the samples from Britain and this sample from France. STRUCTURE results were not as easy to interpret as the DAPC results. The mean L(K) statistic seemed to suggest that K becomes more informative up to the maximum number of populations assumed; in comparison, the delta K statistic indicated that the optimum number of ancestral populations was K=3, with some support for K4 and K5 (Fig. 26). Therefore, to explore structure at different levels of K, data for K3-K5 were plotted (Fig. 27). The results for K3 showed that both populations from southern Portugal were dominated by one genetic cluster, infrequently found in other populations. This was also supported by K4 and K5. At K5 the STRUCTURE results supported the DAPC, in which one genetic cluster is predominantly only found in the populations from western Ireland.

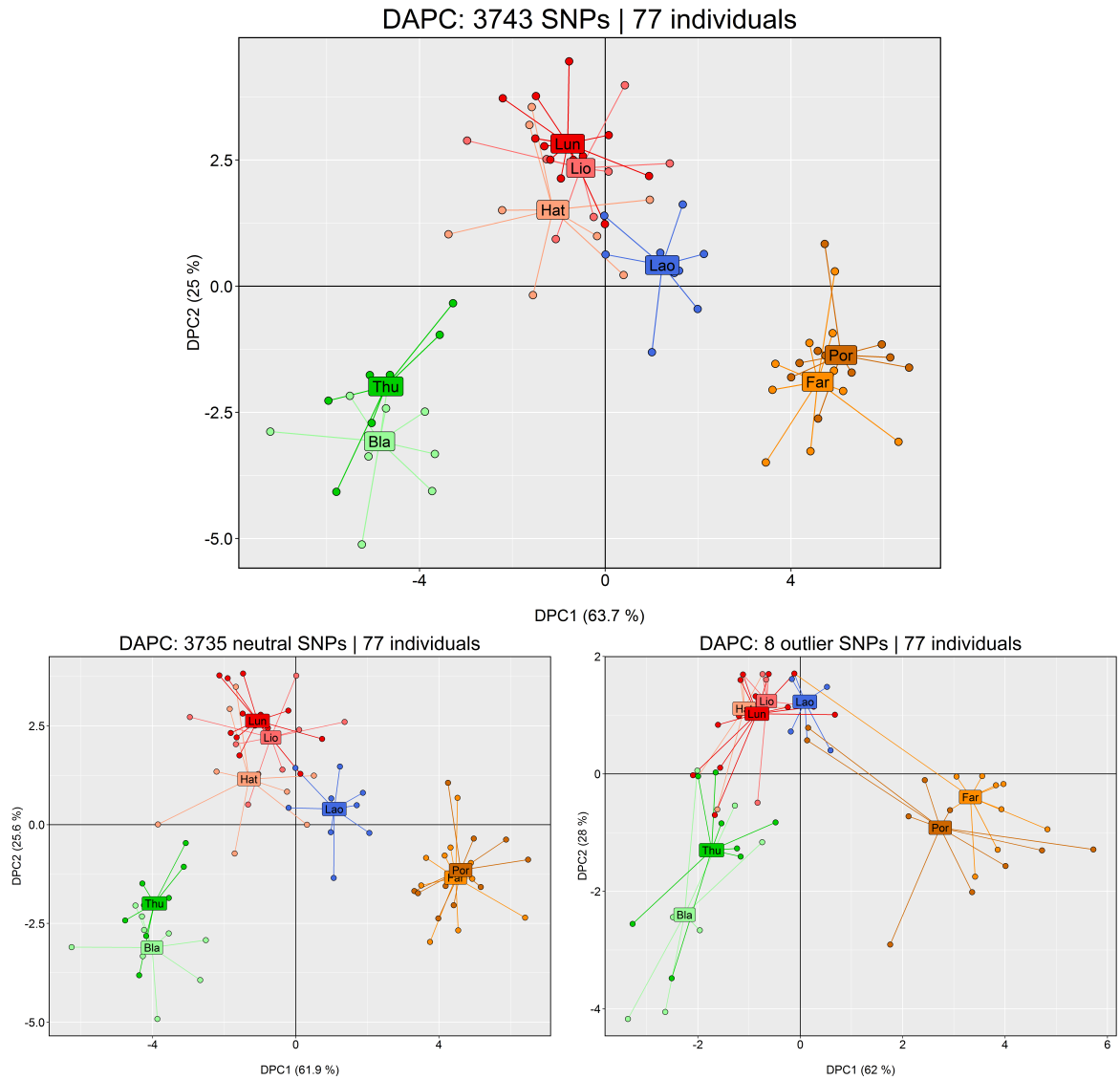


Figure 25: Discriminant analysis of principle components (DAPC) using all SNPs (top), putatively neutral SNPs (left) and outlier SNPs (right). For each DAPC, points represent individuals and colours denote the sampling region of origin (Britain, shades of red-pink; France, blue; Ireland, shades of green; Portugal, shades of orange-brown).

4.3.5 Outlier SNPs

Across all four outlier tests, 131 loci were identified as outliers but only eight of these loci were identified in two or more methods (Fig. 28). The consensus sequences (150 bp) containing these eight loci were submitted to blastx (NCBI) to check whether they matched any translated proteins. One of these eight consensus sequences (locus 1813) had several hits to uncharacterised or hypothetical proteins from scleractinian corals and a sea cucumber (*Apostichopus japonicus*).

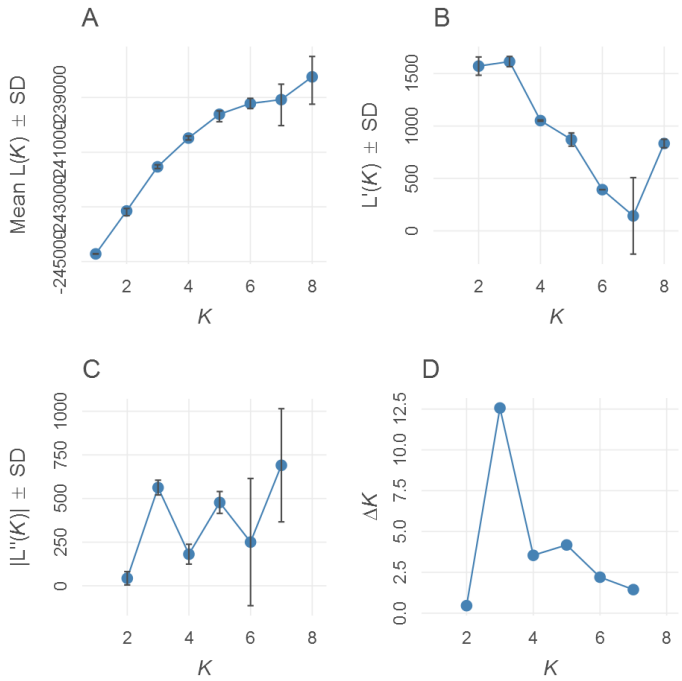


Figure 26: Interpreting *K* using the L(*K*) and the delta *K* methods. (A) mean L(*K*), (B) L′(*K*), (C) L″(*K*), (D) delta *K*.

4.4 Discussion

The results of this study demonstrate that isolating and genotyping genome-wide SNPs in *E. verrucosa* using a RRS approach was successful, despite the constraints in obtaining clean, high molecular weight DNA from many samples. Quality control steps showed that some individuals had a lot of missing data, which were subsequently removed as they could have biased downstream analyses. Population structure analyses indicated that SNP markers showed very similar broad patterns of spatial genetic structure to microsatellite markers; however, slight differences were observed in the inbreeding coefficient calculated between the two types of markers. This SNP study provides the first insight into the population genomics of *E. verrucosa*, a species of major conservation importance in the UK and in the wider northeast Atlantic Ocean.

4.4.1 Feasibility of nextRAD sequencing

In *E. verrucosa*, isolating SNPs from across the genome appears to be feasible using nextRAD sequencing, even with the low-quality DNA obtained from many of our samples. Indeed, a number of samples had to be removed due to missing data, which may have resulted from poor DNA quality. Fragmented DNA (or mutations) can result in the loss of sites where the nextRAD selective primers are designed to

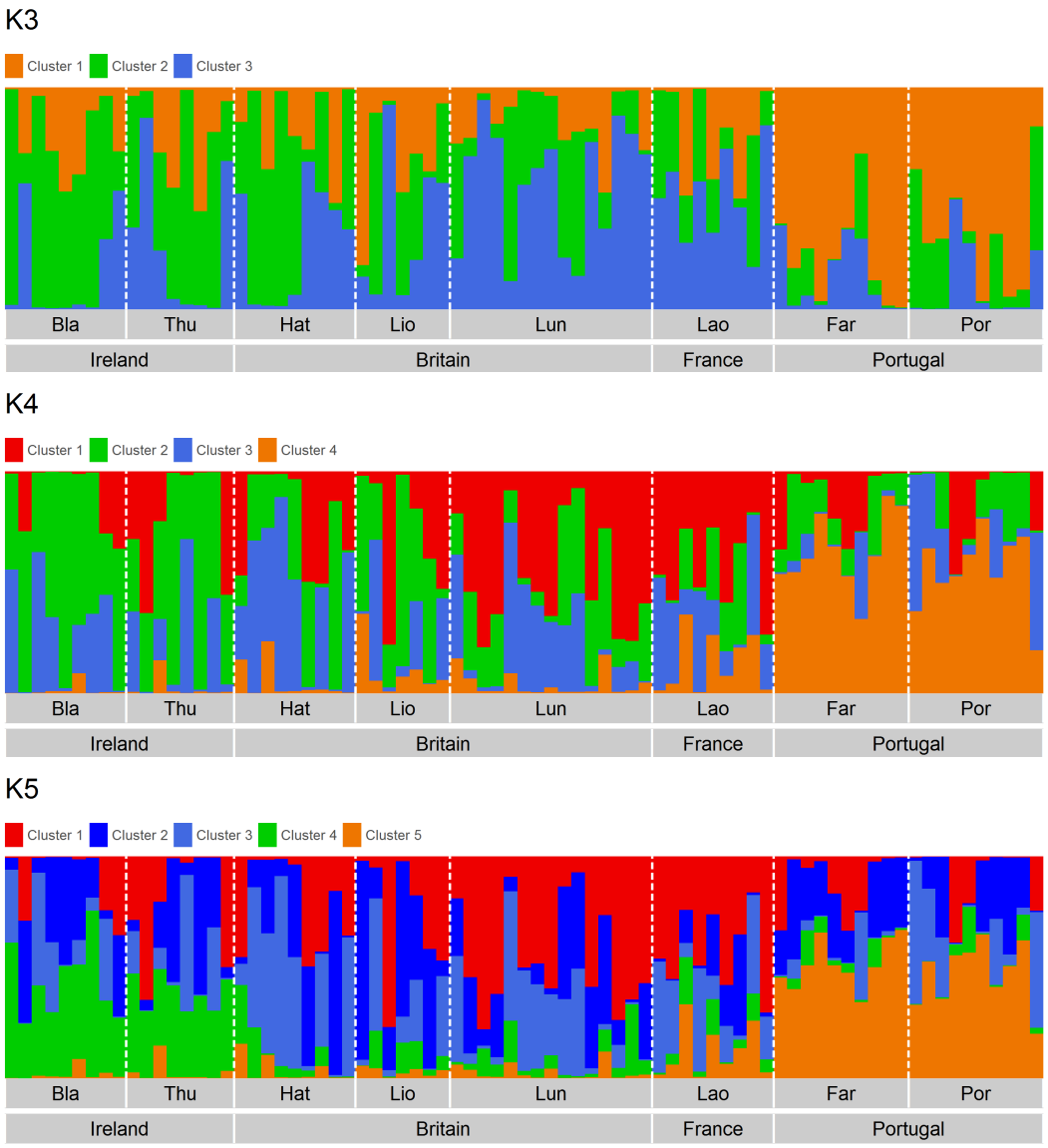


Figure 27: STRUCTURE barplots showing individual membership proportions to *K* ancestral populations (clusters). Each bar represents an individual and colours denote membership proportions to each cluster. Individuals are organised by sampling site and the country of origin.

bind to (or restriction enzyme cut-sites in RADseq), potentially leading to allelic dropout. Allelic dropout from these approaches can introduce bias into the dataset because it may produce overestimates of genetic variation within and between populations (Gautier et al. 2013). However, in this study, although most *E. verrucosa* DNA samples showed evidence of degradation (Appendix A5), the samples retained after quality control contained relatively low amounts of missing data (<20 %, Fig. 21). In RADseq and other methods that utilise NGS technology, missing data have been found to be quite prevalent regardless of the species studied (Flanagan et al. 2017). The way in which this missing data is dealt with can

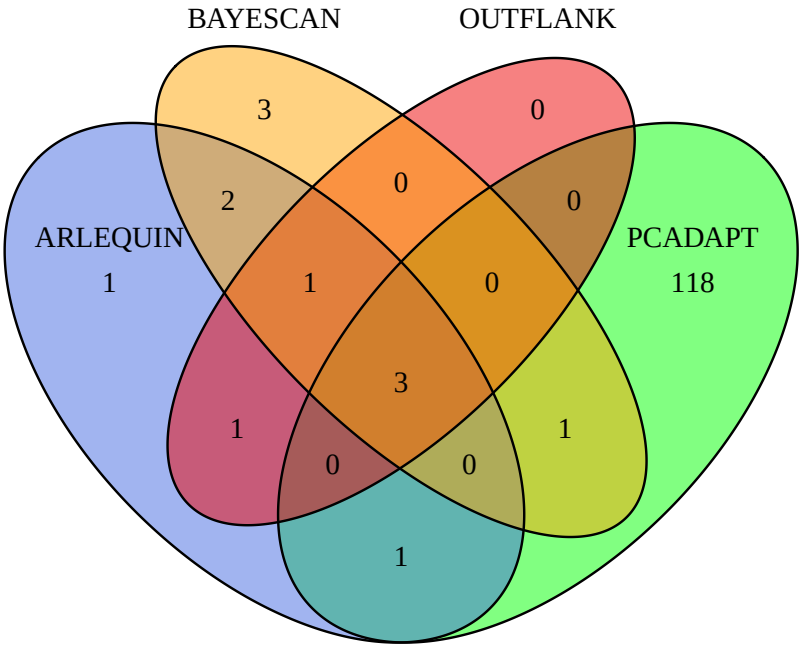


Figure 28: Venn diagram of the outlier SNPs detected by four differentiation-based methods: Arlequin, Bayescan, OutFLANK and PCAdapt.

have consequences for downstream analyses because, if not enough missing data is discarded, then one can run the risk of detecting spurious (non-biological) signals (Flanagan et al. 2017). On the other hand, discarding too much missing data may introduce its own biases; for example, Huang & Knowles (2016) suggested that as the tolerance for missing data becomes more stringent, the mutational spectrum represented in the sampled loci is reduced, leading to the exclusion of loci with the highest mutation rates. Nonetheless, it has been shown that informative SNP datasets can still be acquired when DNA is degraded (Graham et al. 2015) and where there is a high proportion of missing data (Chattopadhyay et al. 2014; Hodel et al. 2017; Tripp et al. 2017).

4.4.2 Microsatellites vs SNPs

The microsatellite and SNP markers showed almost identical patterns of spatial structure. Both sets of markers showed three distinct genetic clusters organised into samples from Britain-France, southern Portugal and western Ireland, with some weak genetic differentiation between samples from southwest Britain and northwest France. The latter result could be due to reduced genetic connectivity across the English Channel, which would imply that drift is the dominant process driving these weak genetic differences. In contrast to *E. verrucosa*, a similar octocoral species, dead man's fingers (*Alcyonium digitatum*), exhibited high genetic

connectivity across this spatial scale (Holland et al. 2017), which suggests that reduced connectivity across the English Channel in *E. verrucosa* may be related to life history strategies rather than a hydrological barrier to dispersal. Although the effects of mid-channel currents and near-shore eddies (Dauvin 2012) on cross-Channel larval dispersal remains to be explored, previous research has also identified a potential genetic break around Brittany in some taxa, including polychaetes (Jolly et al. 2005), nematodes (Wielgoss et al. 2008) and bivalves (Becquet et al. 2012). Conversely, other taxa show panmixia or high connectivity across this area, such as shrimps (Luttikhuizen et al. 2008), lobsters (Triantafyllidis et al. 2005), sea stars (Baus et al. 2005) and cuttlefish (Wolfram et al. 2006). This supports the notion that migration across the English Channel is likely dependent on species-specific factors rather than a universal barrier to gene flow.

What is also apparent from both marker systems is that within each region (i.e. southwest Britain, northwest France, southern Portugal and northwest Ireland), sampling sites are connected by high gene flow and/or have large N_e. For example, in the microsatellite study, little differentiation was observed between the two most distant sites within Britain (Sawtooth Ledges in Lyme Bay and Skomer Island), implying that gene flow can potentially occur up to distances of 470 km. This could not be tested in the SNP study because these sites were not included in the nextRAD libraries, but based on the genetic similarity of Hat, Lio (both Isles of Scilly sites) and Lun (Lundy Island), this provides evidence for gene flow potentially up to spatial scales of 192 km (Hat and Lun). Moreover, the single sampling site from France, Lao, showed some genetic similarity with all three sites from Britain, suggesting that some gene flow may occur at scales up to 447 km. However, as mentioned previously, this similarity could also be explained by large Ne which can mitigate the influence of drift; therefore, further work is required to assess whether the English Channel does indeed constitute a partial barrier to gene flow in *E. verrucosa*.

The three distinct genetic groups found using both marker systems is robust evidence that the genetic structure of *E. verrucosa* across the sampled area has been adequately resolved. Still, the precise drivers of some of these patterns are yet to be determined with confidence. In the microsatellite study, IBD appeared to explain the differentiation observed between sampling sites from Britain, France and Portugal (Holland et al. 2017). Although the SNP study contained many fewer sampling sites, almost identical genetic patterns were found using differentiation indices (i.e. F_{st} and D) and DAPC, which supports this explanation. However, it is still unclear whether the genetically distinct colonies in northwest Ireland are driven by neutral processes (i.e. barrier to gene flow) or adaptive processes (i.e. selection).

By isolating genome-wide SNPs using NGS, this allowed the exploration of outlier SNPs that may be potentially under divergent selection. The DAPC patterns with and without the eight outlier SNPs detected were generally comparable to the DAPC using all SNP loci (Fig. 25). This suggests that both putatively neutral and putatively adaptive SNPs may be contributing to the genetic distinctiveness of western Ireland colonies. Interestingly, one of these SNP loci matched multiple translated proteins from stony corals on the NCBI database, but the functions of these proteins are currently unknown. Ultimately, to fully explore the potential role of selection in these colonies inhabiting the peripheral range of *E. verrucosa*, more genomic resources and gene annotations are required for this species or a closely related species.

In comparison to the spatial patterns of genetic structure, measures of genetic diversity were slightly contrasting between microsatellite and SNP markers. Where microsatellite markers found a mixture of significant and non-significant heterozygote deficiencies across sites (Table 1 in Holland et al. 2017), SNP markers found significant heterozygote deficiencies across all sites included in the nextRAD study (Table 7). In the microsatellite study, both Lio and Lun showed non-significant deficiencies, whereas they showed significant deficiencies in the SNP study. A similar pattern was found for Bla; however, all other sites from the SNP study (Thu, Lao, Far and Por) agreed with the microsatellite markers. This discrepancy in some sites is difficult to explain and could be due to one or more factors that are known to cause heterozygotes deficiencies: (i) the locus is under selection; (ii) the presence of null alleles that may lead to an excess of homozygotes; (iii) inbreeding; and (iv) the presence of population substructure leading to the Wahlund effect. The latter two are the most likely explanations for the sites showing heterozygote deficiencies in both marker systems, which may be the product of low dispersal capacity of larvae and high self-recruitment at these sites. Selection acting on certain loci could explain the discrepancies between the two marker systems, particularly as some loci in both the microsatellite and SNP studies were identified as being potentially under divergent selection (or genetic hitchhiking). However, in the SNP study, genetic diversity statistics were run without the eight outlier loci (Table 7) and significant homozygote excesses were still present, which indicates that selection is unlikely to be driving this deficiency of heterozygotes. Of course, it is possible that there were false-negatives in the outlier tests of both studies, meaning some loci were considered neutral when they are actually under selection. Null alleles were controlled for in both microsatellite and SNP studies; however, it is also possible that false-positives were present here which may have caused the discrepancy between the two marker systems.

Compared to other octocoral species, *E. verrucosa* appears to have lower

genetic diversity when both microsatellites and SNPs are considered (Table 8). For example, the closely related species *Eunicella cavolini* (Masmoudi et al. 2016) and *E. singularis* (Costantini et al. 2016), *A. digitatum* (Holland et al. 2017), as well as two other Mediterranean octocorals, *Corallium rubrum* (Ledoux et al. 2010) and *Paramuricea clavata* (Mokhtar-Jamai et al. 2011) all have higher mean H_e and allelic richness at microsatellite markers, and the Pacific deep-sea octocoral *Swiftia simplex* has higher mean H_e at SNP markers (Everett et al. 2016). Although low genetic diversity observed from the 13 microsatellite markers may be explained by low polymorphism at some loci (Holland et al. 2017), the precise causes of this overall low genetic diversity remain to be determined but are perhaps the result of a combination of processes such as site-specific inbreeding, historic bottlenecks and the purging of alleles by strong drift or selection.

4.4.3 Implications for conservation and MPAs

Eunicella verrucosa is a conservation priority in England and Wales and is internationally recognised as a species facing a very high risk of extinction in the medium-term future. Akin to tranche one, several of the MCZs recently designated in tranche two (January 2016) around southwest Britain (e.g. Bideford to Foreland Point, Hartland Point to Tintagel, and Runnel Stone) specifically identify *E. verrucosa* as a protected feature in their designation listing. In addition to protection from MCZs, it has been found that 60% of *E. verrucosa* colonies recorded by diver surveys in southwest Britain fall within areas protected by various other pieces of EU legislation (Pikesley et al. 2016). However, not all of these MPAs prohibit bottom trawling (e.g. The Manacles, Whitsand and Looe Bay, and Chesil Beach and Stennis Ledges MCZs), the absence of which has been shown to positively affect the ability of pink sea fan populations to recover, albeit over a long time period of up to 20 years (Kaiser et al. 2018). This suggests that a large proportion of *E. verrucosa* in Britain remain vulnerable to anthropogenic disturbance and that the current level of protection afforded by MCZs is in some areas insufficient (Lieberknecht & Jones 2016; Pikesley et al. 2016).

The genetic data for *E. verrucosa* presented here (using both microsatellites and SNPs) highlight interesting findings relevant to the conservation of this ecologically important sessile species at local and regional scales; this has implications for both single-site feature designations and network connectivity. For instance, the microsatellite and SNP results suggest there is high genetic connectivity among populations of *E. verrucosa* around coastal areas of southwest Britain. This may suggest high genetic connectivity, but large N_e can also produce

Table 8: Comparison of genetic diversity in temperate octocorals using microsatellite and SNP markers.

Family Species	Sea	Marker	Sites; *N*	No. loci	mean H_e	mean A_r	Reference
Alcyoniidae							
Alcyonium digitatum	Atl	Msat	20; 648	8	0.63	4.18	Holland et al. 2017
Coralliidae							
Corallium rubrum	Med	Msat	40; 1,222	10	0.74	7.30	Ledoux et al. 2010
Gorgoniidae							
Eunicella cavolini	Med	Msat	19; 584	7	0.56	4.24	Masmoudi et al. 2016
Eunicella singularis	Med	Msat	13; 301	6	0.53	3.58	Costantini et al. 2016
Eunicella verrucosa	Atl	Msat	27; 905	13	0.42	2.58	Holland et al. 2017
	Atl	SNP	8; 77	3,743	0.25	n/a	This study
Plexauridae							
Paramuricea clavata	Med	Msat	39; 1114	6	0.74	6.48	Mokhtar-Jamai et al. 2011
Swiftia simplex	Pac	SNP	4; 23	786	0.26	n/a	Everett et al. 2016

Atl, Atlantic; Med, Mediterranean; Pac, Pacific.
Msat, microsatellite; SNP, single nucleotide polymorphism.
N, number of individuals genotyped; H_e, expected heterozygosity; A_r, allelic richness.
n/a, statistic not relevant because biallelic SNPs were used.

similar patterns. N_e is difficult is estimate in marine species (Hare et al. 2011), although methods to assess N_e based on linkage disequilibrium and coalescence analysis are being refined and developed that also incorporate genomic data (Nunziata & Weisrock 2018; Marandel et al. 2018), which may improve estimates of N_e for marine species. Moreover, inferences of gene flow suggested that populations of pink sea fan in southwest Britain act as a source for adjacent populations across the English Channel, highlighting the value in protecting these populations. In the UK, the current recommendation for maintaining ecological connectivity between discrete habitats is the positioning of a MPA every 80 km or less (Roberts et al. 2010). Considering our findings from this population genetics study, it appears that these distances between MPAs would generally be sufficient to maintain genetic connectivity in *E. verrucosa* across southwest England and Wales. Of course, this assumes that contemporary local oceanic currents are able to facilitate the transport of enough larvae between sites, whether by a continuing stepping-stone process or a single dispersal event.

The genetic distinctiveness of *E. verrucosa* populations from Donegal Bay in northwest Ireland, detected with both marker systems in this study, reinforces an argument for protecting these local sites. Marginal populations often contain rare alleles (the highest extent of private alleles were found at these sites in the microsatellite study), but may recruit more slowly and be demographically isolated, implying reduced resilience to disturbance and therefore an increased need for protection. Consequently, although *E. verrucosa* are not specifically protected in the Republic of Ireland, the development and effective management of the Donegal Bay SAC could be crucial to the persistence of the distinct genetic variants of *E. verrucosa* found in this area.

A similar pattern was found for pink sea fans in southern Portugal, which are also genetically distinct from other regions, but this is likely driven by the large geographical distance that separates these populations from those north of the Bay of Biscay. This suggests that populations of *E. verrucosa* would benefit from protection across the species range as a connected meta-population. However, the implementation of such a measure for a single species is unlikely given the dwindling resources available to conservation managers. Nevertheless, highlighting these species-specific areas of unique diversity for consideration may prove valuable if national governments move towards an ecosystem-based management approach, whereby whole ecosystems are protected from disturbance.

4.4.4 Knowledge gaps and future research

The two marker systems (microsatellites and SNPs) discussed in this chapter showed very similar results across the sites studied, which means we have a relatively robust understanding of the genetic patterns across the range sampled in *E. verrucosa*. However, increasing the number of genotyped individuals and sites for the SNP study would undoubtedly help to resolve whether there is any fine-scale genetic structure within regions, particularly for sites that were included in the microsatellite but not the SNP study. Moreover, both marker systems only had samples from locations across the middle and northern range of *E. verrucosa* – no areas of the southern range were included in these studies (i.e. the Mediterranean). In addition, there were several locations in the middle and northern range that could not be sampled due to logistical and time constraints (e.g. western and southern Ireland, and northern Spain). Pink sea fans are apparently found in sufficient numbers for meaningful population genetic analysis in Tarragona (southeast Spain, Mediterranean), Marseille (southern France, Mediterranean), northern Spain (Aurelle et al. 2017), and Galway Bay (western Ireland, Wood 2013). Inclusion of samples from these locations would enable three main hypotheses to be tested in future research:

Firstly, the role of IBD in driving population structure between *E. verrucosa* colonies from southwest Britain, the Bay of Biscay and southern Portugal could be further explored by sampling colonies from northern Spain (Fig. 29). The null hypothesis would state that IBD is the main driver of population structure across these locations, which would agree with the results from this chapter and Holland et al. (2017). The alternative hypothesis would suggest that colonies from northern Spain are genetically distinct from other locations, perhaps due to a barrier to gene flow or local adaptation, but that IBD still explains much of the differentiation observed between colonies from southwest Britain, northwest France and southern Portugal.

Secondly, additional samples from western or southern Ireland would allow further exploration of the relative contribution of gene flow, drift and selection to driving the distinct genetic profiles observed in Donegal Bay (northwest Ireland). The null hypothesis would suggest that IBD is the primary driver of this differentiation, which would support a stepping-stone model of dispersal across southwest Britain and Ireland. The first alternative hypothesis would suggest that IBD is not a primary driver, but that other neutral processes have mainly contributed to this differentiation (e.g. barrier to gene flow, past bottlenecks or periglacial refugia). In contrast, the second alternative hypothesis would propose

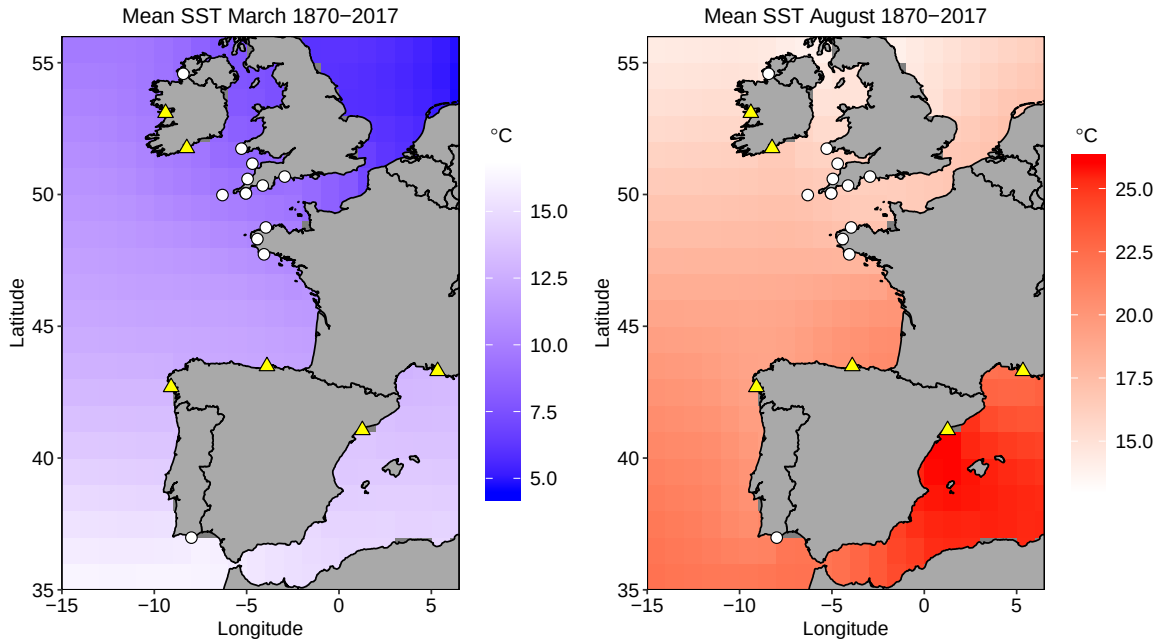


Figure 29: Map of mean minimum (March) and mean maximum (August) sea surface temperature (SST) from 1870-2017 across the northeast Atlantic and the Mediterranean. White dots represent sampling sites from Holland et al. 2017 and the SNP study presented in this chapter. Yellow triangles represent sites in which obtaining samples of *E. verrucosa* would be advantageous for further exploring the drivers of population genetic structure in future research. SST data was extracted from the Met Office Hadley Centre Sea Ice and Sea Surface Temperature dataset.

that natural selection is driving this differentiation, whereby these *E. verrucosa* colonies have become adapted to specific environmental conditions in Donegal Bay. In fact, this area represents the coldest local sea temperatures across the entire range of *E. verrucosa* (Fig. 29), so it may be possible that these colonies have developed a tolerance to these sea temperatures, and that one or more of the loci studied here are involved in this adaptive tolerance. Indeed, Pivotto et al. (2015) demonstrated that thermo-tolerance varied greatly along a depth gradient in a closely related species, *E. cavolini*, which implies that some octocoral species have the capacity to adapt to different thermal regimes.

Lastly, samples from the Mediterranean would enable a direct comparison of the genetic diversity and population structure from sites in the southern range with sites in the middle and northern range of *E. verrucosa*. A number of questions could be put forward here, for example: (i) are spatial patterns of genetic diversity and population structure comparable between the two basins?; (ii) how connected are populations from the northeast Atlantic and the Mediterranean?; and (iii) are colonies inhabiting the periphery of the southern range locally adapted to the

warmer temperatures? Exploring this question of local adaptation is of particular interest because of the vast differences in sea temperatures present at the range limits of *E. verrucosa*; for example, from 1870-2017, the average lowest temperature at Donegal Bay was 8.2°C (in March), whereas the average highest temperature at Tarragona was 28.6°C (in August) (Fig. 29). Furthermore, investigating spatial patterns of population genetic structure across environmental gradients has enabled previous studies to reveal how different temperature regimes contribute to shaping the genetic structure of many marine organisms (Benestan et al. 2016b; Diopere et al. 2017; Van Wyngaarden et al. 2018; Lehnert et al. 2018).

In conclusion, this chapter has confirmed that nextRAD sequencing is an appropriate RRS method for genome-wide SNP discovery in *E. verrucosa*. However, this was only possible after modifying a salting-out protocol which optimised the amount of high molecular weight DNA extracted. Both markers systems, 13 microsatellites and 3,743 SNPs, showed almost identical patterns of spatial genetic structure, but slightly differing patterns were observed in the inbreeding coefficient at some sites. Results from both marker systems suggest that the current network of MPAs in southwest England and Wales is sufficient for maintaining genetic connectivity in this species. Inbreeding was found to be site-specific, but whether this has an impact on fitness and the long-term resilience of the sites in question is currently unknown. Although we have a robust understanding of the spatial genetic patterns at the sites sampled, collecting samples from areas that were not represented previously will be advantageous going forward as it will enable a more thorough and complete assessment of these patterns across the entire range of *E. verrucosa*. Considering the ecological importance of this species, acquiring this information will be vital for the international conservation of this vulnerable species and for monitoring how climate change may impact populations at the range edges.

Chapter 5: SNP discovery in European lobster (*Homarus gammarus*) using RAD sequencing

---

This chapter is based on a paper published in the journal *Conservation Genetics Resources*. The reference is given below and the full paper is available in the Appendix.

Jenkins TL, Ellis CD, Stevens JR (2018) SNP discovery in European lobster (*Homarus gammarus*) using RAD sequencing. *Conservation Genetics Resources* https://doi.org/10.1007/s12686-018-1001-8.

K. Moore (Exeter Sequencing Service) prepared the RADseq libraries.

5.0 Abstract

The European lobster (*Homarus gammarus*) is a decapod crustacean with a high market value and therefore their fisheries are of major importance to the economies they support. However, over-exploitation has led to profound stock declines in some regions such as Scandinavia and the Mediterranean. To manage this resource sustainably, knowledge of population structure and dispersal is crucial to inform management about stock structure, connectivity and food traceability. In this chapter, restriction-site associated DNA sequencing was used to develop novel SNP markers from 55 individuals originating from 27 geographically separate sampling locations which encompassed much of the species' range; SNPs were quality filtered, ranked using differentiation statistics and the top 96 SNPs adequate for primer design were retained. SNP markers were developed with the aim of maximising the power to detect genetic differentiation between: (i) Atlantic and Mediterranean lobsters and (ii) Atlantic lobsters. This SNP panel provides a useful resource for future studies of population genetic structure and assignment in *H. gammarus*.

5.1 Introduction

The European lobster, *Homarus gammarus* (Fig. 30), is a large decapod crustacean belonging to the family Nephropidae. They are found in most coastal seas of the northeast Atlantic, historically ranging from northern Morocco to northern Norway, including the British Isles and Skagerrak, and the Mediterranean and western parts of the Black Sea. Though, they are extremely rare across their southern range (i.e. Morocco, parts of the Mediterranean and the Black Sea). The species' high market value makes it a highly-prized seafood product, so its fisheries are of great importance to the local and regional economies they support. However, current and historical over-exploitation has led to stock declines, some of which have been quite profound in several regions (e.g. Scandinavia, Mediterranean) and from which recovery has been slow or stagnant (Kleiven et al. 2012). This has led to the rearing of *H. gammarus* larvae in lobster hatcheries to produce juveniles which are released into the wild to supplement productive stocks where the risk of over-exploitation is high (Ellis et al. 2015c). To manage this resource sustainably, developing genetic resources that enable the assessment of population structure and connectivity, as well as the ability to assign lobsters to their site of origin, is crucial to inform management about dispersal, stock boundaries and connectivity, and food traceability.

Over the last decade, genetic diversity and population structure has been

Figure 30: European lobster (*Homarus gammarus*).

investigated in *H. gammarus* using traditional molecular markers including random amplification of polymorphic DNA (RAPDs) (Ulrich et al. 2001), allozymes (Jørstad et al. 2005), mtDNA restriction fragment length polymorphisms (RFLPs) (Triantafyllidis et al. 2005) and microsatellites (Huserbraten et al. 2013; Watson et al. 2016; Ellis et al. 2017). However, concerns over low sample sizes and geographical coverage, and limitations associated with the molecular markers used in these previous studies, question the power of these studies to adequately resolve the underlying population genetic structure in this species. The isolation of thousands of genome-wide SNPs has commonly been attributed to maximising the power to resolve spatial patterns of genetic variation, showing particular promise for detecting subtle population structure in highly dispersive marine species that exhibit typically weak genetic differentiation (e.g. American lobsters, Benestan et al. 2015; great scallops, Vendrami et al. 2017).

The aim of this study was to isolate thousands of genome-wide SNPs using RADseq and develop a small panel containing the most informative SNP markers that captures weak genetic differentiation between our sampling sites. SNPs were chosen with the aim of maximising the power to detect genetic differentiation (i) between Atlantic and Mediterranean lobsters and (ii) between Atlantic lobsters. Previous studies have suggested that *H. gammarus* from the Atlantic and the Mediterranean are two distinct groups (e.g. Jørstad et al. 2005; Triantafyllidis et al. 2005); thus, a SNP panel that can accurately assign unknown individuals to the Atlantic Ocean or the Mediterranean Sea may be a useful tool for food traceability and the monitoring of resources.

5.2 Materials and methods

5.2.1 Sample collection

Tissue samples of adult European lobsters were mostly obtained by establishing national and international collaborations with a diverse number of personnal across western Europe. The most common method used to collect samples comprised sending a sampling kit containing sterile gloves, 2 ml tubes filled with 95-100 % ethanol and a signed return package to a sampler who had agreed to collect tissue samples. The sampler then took a tissue sample from each lobster and placed the sample into one of the tubes provided. When sampling was complete, the sampler sent the samples back to the lab, placing a note inside specifying the geographical origin of sampled lobsters. Samplers consisted of fishermen, shellfish merchants / suppliers, restaurant owners, hatchery technicians, marine institutions, governmental bodies (e.g. IFCAs, BIM, etc.) and scientific researchers. Some tissue samples were also taken by myself or archived tissue samples were provided by collaborators. The type of tissue collected depended on the sampler and whether the lobster was caught to be sold or for scientific monitoring research. For the majority of samples, a 1-2 cm section from one or two pleopods (swimmerets) (Fig. 31) was removed from each lobster – all lobsters destined to be sold were sampled in this way to maintain optimal condition and to avoid reducing sell-on prices. The remaining tissue samples were taken from either the uropod (v-notches), the pereiopods (walking leg) or the antennae (Fig. 31). All tissue samples were placed in 95-100 % ethanol and stored in a 4°C cold room for long-term preservation.

5.2.2 DNA extraction

Genomic DNA was extracted from all tissue types using a modified salting-out protocol designed to extract DNA from crayfish exoskeleton (Li et al. 2011). Extracting DNA from pleopod material is challenging because their tough chitin exoskeleton resists grinding and lysing and they contain copious amounts of astaxanthin and other impurities which are not trivial to precipitate and can affect DNA purity (Li et al. 2011). These constraints were addressed by firstly pulverising the tissue using a microbead and a TissueLyser (Qiagen) to homogenise the sample. Secondly, sodium dodecyl sulphate (SDS) was added to the lysis buffer during the digestion to help reduce protein disulphide bonds and denature proteins. Finally, two rounds of ammonium acetate treatment were carried out to completely remove proteins and cellular debris. DNA was precipitated using 100 % cold

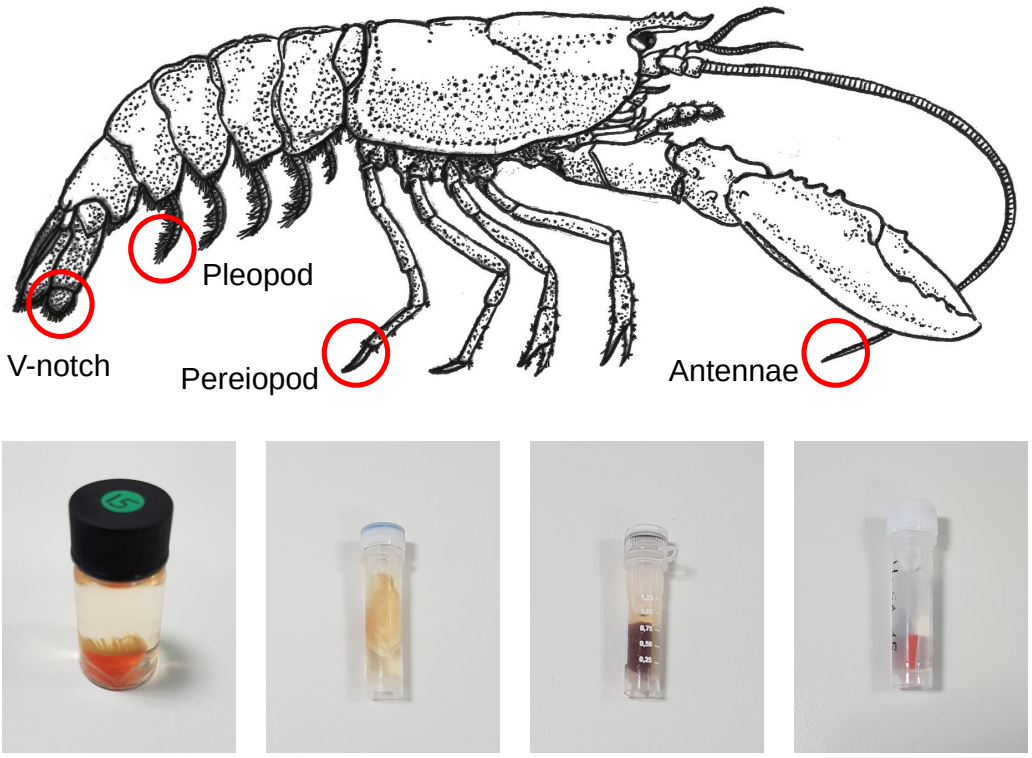


Figure 31: European lobster side-view (top) and tubes containing tissue samples (bottom). Four different tissue types were taken for genetic analysis which was dependent on the individual conducting the sampling: v-notches, pleopods, pereiopods and antennae.

isopropanol and washed with 70 % ethanol, followed by rehydration with 100 μl of nuclease-free water. See Appendix A4 for a detailed step-by-step protocol. The concentration and purity of all DNA extractions were quantified by spectrophotometry using a NanoDrop 1000. In addition, the quality of the DNA samples were further evaluated by running the DNA on a 1 % agarose gel and by quantifying their concentration with fluorometry using the Invitrogen Qubit Assay kit, which measures the amount of double-stranded DNA in the sample.

5.2.3 RAD sequencing

RAD libraries were prepared by the Exeter Sequencing Service using Illumina Nextera XT barcodes. Genomic DNA (400 ng) from each sample was sheared to an average size of 1000 bp using a Covaris E220 sonicator, having previously optimised for a size range of 800-1000 bp. Fragmentation of eight samples was checked using a DNA 1000 screentape (Agilent). The NEBNext Ultra II DNA library preparation kit was used for end-repair, A-tailing and for ligating P2 adapters, and the reactions were purified using AMPure XP magnetic beads (Beckmann Coulter). P2-adapted DNA was then digested using the restriction enzyme SbfI at 37°C for 4

hours and purified using AMPure XP beads to avoid heat-denaturing the enzyme, which can lead to a bias in the libraries. Phased P1 adapters were ligated to the digested fragments and unligated adapters were removed using AMPure XP beads. The P1 adapter was biotinylated at the 5'end of the top strand to enable capture with streptavidin beads. After washing away fragments not bound to the P1 adapter, the DNA was amplified by PCR to add Nextera XT multiplexing barcodes and flow cell attachment regions. Library quality and quantity was assessed using DNA screentapes. Equimolar pooling of the libraries was undertaken before size selection of libraries averaging 660 bp (inserts $\sim$530 bp). The size selected pool was quantified by qPCR and stored at -20°C prior to sequencing.

A pilot study using seven samples of *H. gammarus* was firstly carried out to ensure RAD sequencing using SbfI was appropriate for this non-model species. The SbfI restriction enzyme was initially chosen for two reasons. Firstly, SbfI has been used in a previous study (Benestan et al. 2015) on a closely related species, the American lobster (*Homarus americanus*), and good quality data were obtained. Secondly, a theoretical prediction of the number of loci (RAD-tags) was estimated using the RAD-tag counter from GenePool (Edinburgh, UK). The GC content was set to 0.40 and the genome size was set to 4,250 mb (Animal Genome Size Database). The RAD-tag counter indicated there were approximately 24,480 cut sites and 48,960 RAD-tags in the genome, which was deemed adequate for this study. The RAD libraries of these seven samples were added to a lane of Illumina HiSeq, with single-end sequencing at 150 bp read lengths. After quality control and filtering, preliminary analysis indicated that the number of reads, the depth of coverage, and the number of SNPs discovered was adequate for this study.

After the successful pilot study, a more comprehensive RAD library was designed. As the end goal was to discover and develop a panel of SNPs, the RAD library composed lobsters representing the geographical spread of our sampling coverage (Fig. 32), which was then sequenced at high coverage. This approach attempted to incorporate as much of the genetic variation potentially present in our geographic sampling at the time of library construction (additional sites were sampled after this RAD library was sequenced). In addition, sequencing at high coverage meant that more reads were merged to form RAD-tags and call SNPs which would theoretically improve the reliability of calling 'true' SNPs. This RAD library was composed of 48 lobsters, with one-three individuals per sampling site (Table 9; Fig. 32). The library was then sequenced on two lanes of an Illumina HiSeq 100 bp paired-end rapid run platform. Paired-end sequencing was chosen because the reverse reads can be aligned back to the forward reads to increase the length of a RAD-tag, which may allow SNPs that are positioned at the end of

the RAD-tag to be considered for primer development.

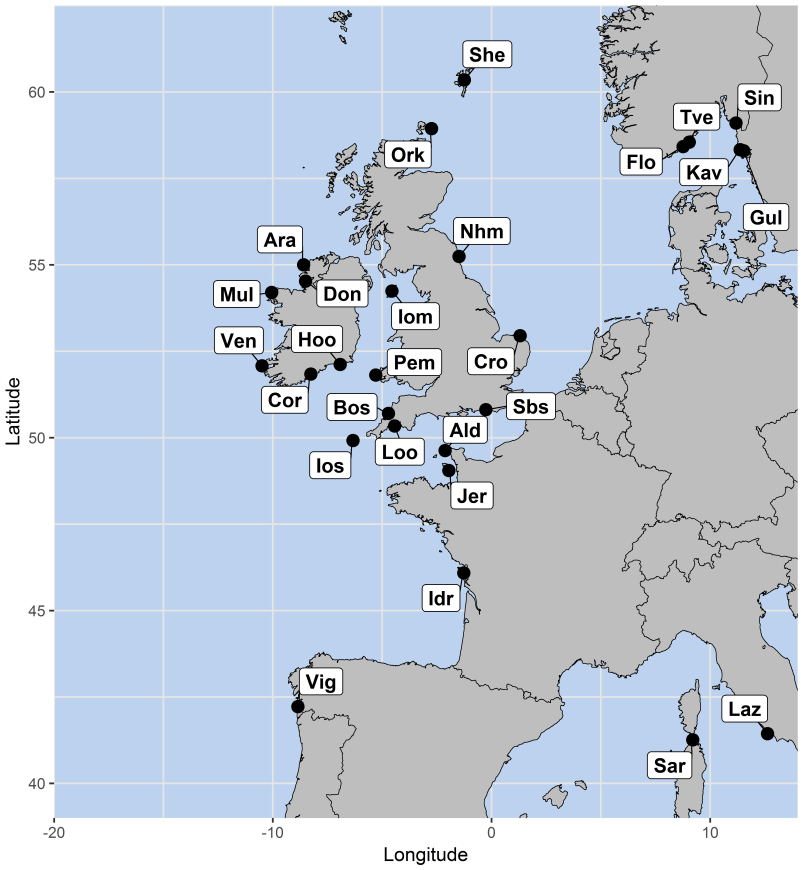


Figure 32: European lobster: RADseq sampling coverage

5.2.4 Quality control

Quality of the raw reads was initially examined using the FastQC software (Babraham Bioinformatics). The sequencing data provided by the Exeter Sequencing Service had already been demultiplexed, which meant that fastq files were received for each individual (two files per individual because of paired-end sequencing). The use of phased adapters meant there were one of four different combinations at the start of each forward read: TGCA, ATGCA, CATGCA or GCATGCA (nucleotides marked in red represent part of the SbfI cut-site). As a result, the raw data was trimmed at the start of each read, such that each read started with TGCA, the SbfI cut-site.

Before further bioinformatics, raw reads from both the pilot study (seven individuals) and the RAD library (48 individuals) were pooled, creating a final RAD dataset containing 55 individuals (Table 9). Raw reads were then cleaned and truncated to 97 bp using the `process_radtags` program in the Stacks software

v1.45 (Catchen et al. 2013). A read was discarded if the score fell below a 90 % probability of being correct (a raw Phred score of 10). Reads were truncated to 97 bp because this was the shortest read length in the dataset and Stacks requires all reads to be the same length. No barcode argument was provided because the data had already been demultiplexed.

5.2.5 Building loci *de novo*

The formation of RAD loci was carried out by running the wrapper script `denovo_map.pl` in Stacks. There are three main components of Stacks to consider when executing the `denovo_map.pl` wrapper script for population genomic analyses: `ustacks`, `cstacks` and `populations`. The `ustacks` program aligns the cleaned reads generated by `process_radtags` into exactly matching stacks (or putative alleles) and then builds loci and calls SNPs *de novo* for each individual. With loci built for each individual, the `cstacks` program attempts to match loci across samples to create a catalog of loci across all of the samples. The `populations` program computes population genetics statistics, as well as exporting SNP genotypes in user-defined formats (e.g. Genepop format). The default assumes each individual in the RAD library is a separate sample; however, the user can submit a population map that specifies the sample (i.e. population or sampling site) that each individual belongs to.

There are three main parameters that control locus formation and SNP calling in Stacks: the minimum number of reads required to form a stack (*m*; `ustacks`), the number of mismatches allowed between stacks to merge them into a locus (*M*; `ustacks`) and the number of mismatches allowed between stacks during the construction of the catalog (*n*; `cstacks`). Exploring these parameters is important for optimising the number of assembled loci (RAD-tags), polymorphic loci and SNPs discovered in RAD datasets (Paris et al. 2017). In this study, following the advice of Paris et al. (2017), a strategy was implemented to optimise these three parameters in our RAD dataset. Each parameter was sequentially changed while keeping the other two constant and the highest number of r80 polymorphic loci was considered the optimum parameter set. The r80 polymorphic loci are identified by adjusting the *r* parameter in the `populations` program, whereby a locus must be present in at least 80 % ($r = 0.80$) of the individuals in a population for it to be processed. For this optimisation process, no population map was specified which meant that a locus had to be present in at least 44 individuals for it to be processed. All other parameters were set to their default values during the optimisation of *m*, *M* and *n*.

Table 9: European lobster sampling information. A total of 55 individuals from 27 sampling sites were used in the preparation of RAD libraries.

Country	Site	Code	*N*	Latitude	Longitude	Tissue type	Year
Britain	Boscastle	Bos	2	50.70	-4.71	Pleopods	2013
	Cromer	Cro	2	52.95	1.31	Pleopods	2016
	Isle of Man	Iom	2	54.24	-4.55	Pleopods	2016
	Isles of Scilly	Ios	2	49.92	-6.33	Pleopods	2016
	Looe Harbour	Loo13	1	50.34	-4.43	Pleopods	2013
		Loo16	1	50.34	-4.43	Pleopods	2016
	Northumberland	Nhm	1	55.24	-1.52	Pleopods	2013
	Orkney	Ork15	1	58.94	-2.74	Pleopods	2015
		Ork16	1	58.94	-2.74	Pleopods	2016
	Pembrokeshire	Pem	2	51.81	-5.29	Pleopods	2016
	Shetland	She14	2	50.81	-1.53	V-notches	2014
	Shoreham-By-Sea	Sbs	2	50.70	-0.26	Pleopods	2016
Channel Islands	Alderney	Ald	2	49.63	-2.12	Pleopods	2016
	Jersey	Jer	2	49.05	-1.27	Pleopods	2016
France	Île de Ré, La Rochelle	Lro	1	46.09	-2.12	V-notches	2013
	Île de Ré, La Rochelle	Idr16	2	46.09	-2.12	V-notches	2016
Ireland	Aran Island	Ara	2	55.00	-8.59	Pleopods	2016
	Cork Harbour	Cor	2	51.84	-8.26	Pleopods	2016
	Donegal Bay	Don	2	54.53	-8.50	V-notches	2016
	Hook Peninsula	Hoo	2	52.12	-6.92	V-notches	2016
	Mullet Peninsula	Mul	2	54.20	-10.05	V-notches	2016
	Ventry	Ven	2	52.08	-10.05	V-notches	2016
Italy	Lazio	Laz	2	41.44	12.55	Antennae	2013
	Sardinia	Sar13	2	40.12	9.01	Antennae	2013
Norway	Flodevigen	Flo15	1	58.42	8.76	Pleopods	2015
		Flo16	2	58.42	8.76	Pleopods	2016
	Tvedestrand	Tve	2	58.55	9.04	Pleopods	2015
Spain	Vigo	Vig13	2	42.49	-8.99	Pleopods	2013
Sweden	Gullmarfjord	Gul	2	58.30	11.53	Pereiopods	2009
	Kavra	Kav	2	58.33	11.37	Pereiopods	2007
	Singlefjord	Sin	2	59.08	11.12	Pereiopods	2009

N, number of individuals used in the RAD library.

Once the optimum values *m*, *M* and *n* were obtained, `denovo_map.pl` was re-run with the maximum number of stacks at a single locus set to two (`-max_locus_stacks` 2, `ustacks`), meaning only loci with two alleles (biallelic loci) were allowed. This parameter was enforced because, for a diploid species, if more than two alleles are allowed during the *de novo* assembly of loci, the locus assembly could contain merged paralogs or additional alleles introduced as a result of sequencing error (Catchen et al. 2013; Mastretta-Yanes et al. 2014). Individuals were dropped from the analysis if their coverage was lower than 15x.

5.2.6 SNP discovery

SNPs were discovered by running the `populations` program with two sets of parameters. The first run was executed using all 55 samples with no population map, a minimum allele frequency set to 5 % (`-min_maf` 0.05), a maximum observed heterozygosity set to 0.5 (`-max_obs_het` 0.5) and *r* was set to 0.8. The rationale for this run was to search for SNPs that can detect genetic differentiation across broad geographical areas (i.e. between the Atlantic and Mediterranean basins). The second run of populations searched for SNPs that could potentially detect genetic differentiation (if any) between geographical locations within the northeast Atlantic. In this run, the Mediterranean and Skagerrak (Norway and Sweden) samples were removed because early genetic analyses indicated differentiation between these samples and all other Atlantic samples (see Results); accordingly, a population map for only the Atlantic samples was submitted and organised by geography (Table 10). Since most sampling sites were only composed of two or three individuals, the *r* parameter was set to 1, meaning a locus had to be present in every individual in that population for it to be processed. Similarly, the *p* parameter was set to nine (the total number of populations in the population map), which meant that a locus had to be present in all populations in the population map to be eligible. All other parameters remained the same as the first run.

5.2.7 Developing a SNP panel

To develop the SNP panel, a set of criteria were implemented to filter out uninformative SNPs, and SNPs that were inadequate for primer design, whilst retaining the most informative SNPs that were suitable for primer design and high-throughput genotyping on a Fluidigm EP1 system (Fig. 33). Firstly, the SNP dataset generated from the `populations` program were ranked and sorted by highest G''_{st} (Meirmans and Hedrick 2011) using the `diff_stats` function from the R package MMOD (Winter 2012). Other differentiation statistics were computed

Table 10: Stacks population map composed of nine putative populations and 40 individuals.

Population	Individuals	*N*
English Channel	Ald3, Ald4, Jer2, Jer3, Sbs2, Sbs3	6
West Ireland	Ara1, Ara2, Don2, Don3, Ven1, Ven3, Mul3, Mul4	8
Southwest England	Bos55, Bos67, Ios5, Ios6, Loo13_31, Loo16_18	6
Southeast Ireland	Cor1, Cor2, Hoo1, Hoo2	4
North Sea	Cro4, Cro5, Nhm8	3
Irish Sea	Iom1, Iom4, Pem12, Pem14	4
Orkney & Shetland	Ork15_6, Ork16_1, She14_3, She14_4	4
France	Lro_4, Idr16_11, Idr16_13	3
Spain	Vig13_1, Vig13_3	2

N, number of individuals.

(G_{st}, Nei & Chesser 1983; *D*, Jost 2008), but all statistics generally produced very similar results. Post-ranking, information for the top 300 SNPs was recorded in a spreadsheet, including the position of the target SNP in the RAD-tag locus, the locus ID (given by Stacks) and the ranking according to highest $G"_{st}$.

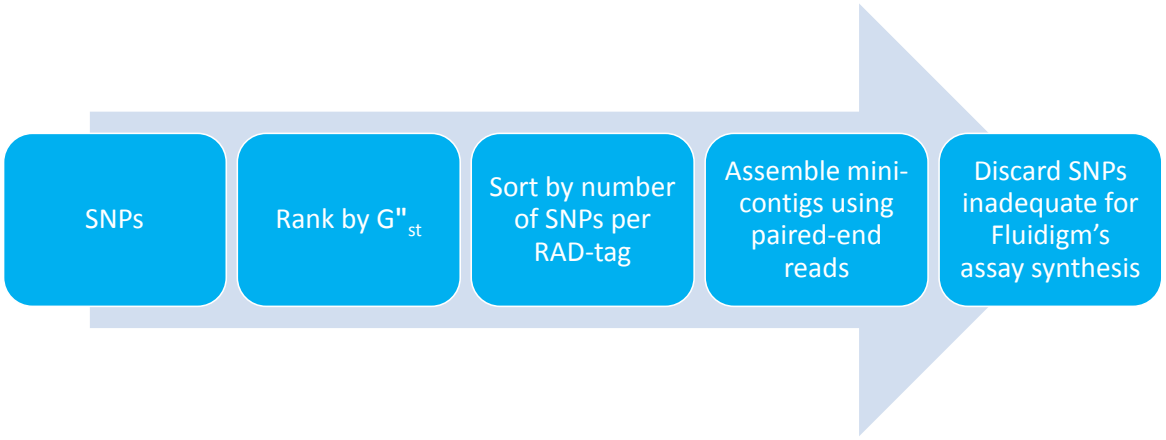


Figure 33: SNP panel development flowchart.

Secondly, the locus ID for each SNP was queried using the Stacks SQL user interface to find out which loci had additional non-target SNPs in the RAD-tag. The SNP database was sorted by the fewest number of SNPs per locus, followed by the highest ranking. The ideal RAD-tag locus contained only one SNP as this was advantageous for primer design; however, RAD-tag loci with two or more SNPs were considered, particularly if they were ranked highly, because high-throughput assay designs, such as the Fluidigm EP1 system, can often deal with non-target SNPs assuming they are at least 30 bp away from the target SNP.

Lastly, to extend the length of RAD-tag loci, mini-contigs were assembled by aligning paired-end reads to the original RAD-tag sequence. First, paired-end sequences were collated for each locus by executing the `sort_read_pairs.pl`

program in Stacks using a whitelist of locus IDs. Then, paired-end reads for each locus were aligned using the `exec_velvet.pl` program with the *M* parameter set to 100, meaning paired-end consensus sequences must be at least 100 bp long. To build mini-contigs, the paired-end consensus sequences were then aligned to the original RAD-tag sequence (97 bp) using the alignment tool in Geneious v.10.1.3 (Kearse et al. 2012). SNPs with flanking sequences $<$25 bp or with non-target SNPs within 30 bp of the target SNP were discarded. Finally, the mini-contig sequences for the remaining SNPs were submitted to Fluidigm's online portal for *in silico* assay design. This performs a quality control analysis on each sequence submitted, checking that SNP assays will be compatible with each other (i.e. no primer-dimers) and that parameters set by Fluidigm are met (e.g. no regions of large repeats, suitable GC content, adequate flanking sequence, etc.).

5.3 Results

5.3.1 Sequencing and quality control

The mean Q scores for the raw sequencing data for all samples ranged from 37-38 with 95 % of reads having $>$Q30 score. The fastQC report indicated that error bars towards the end of the read fell slightly below the threshold of Q28; however, these scores were deemed acceptable and, in any case, the last three bases were trimmed in `process_radtags` during quality control. In total, over 276 million reads (single and paired-end) were generated for the main RAD library (Fig. 34) and a mean average of 97.9 % of reads were retained across all samples after quality control.

5.3.2 Parameter optimisation

Initial results of `denovo_map.pl` using default values of *m*, *M* and *n* indicated high coverage for each RAD locus (38x mean average). Therefore, the parameter *m* was kept at 3 as improving coverage was not an issue and values lower than 3 are not recommended (Paris et al. 2017). The *n* parameter was tested from 2-4 and the highest number of r80 polymorphic loci corresponded to a value of 3 (Fig. 35). Similarly, the *M* parameter was tested from 2-5 and the highest number of r80 polymorphic loci corresponded to a value of 3 (Fig. 35). Subsequently, the optimum set of parameters used to form loci and call SNPs was *m*=3, *M*=3 and *n*=3.

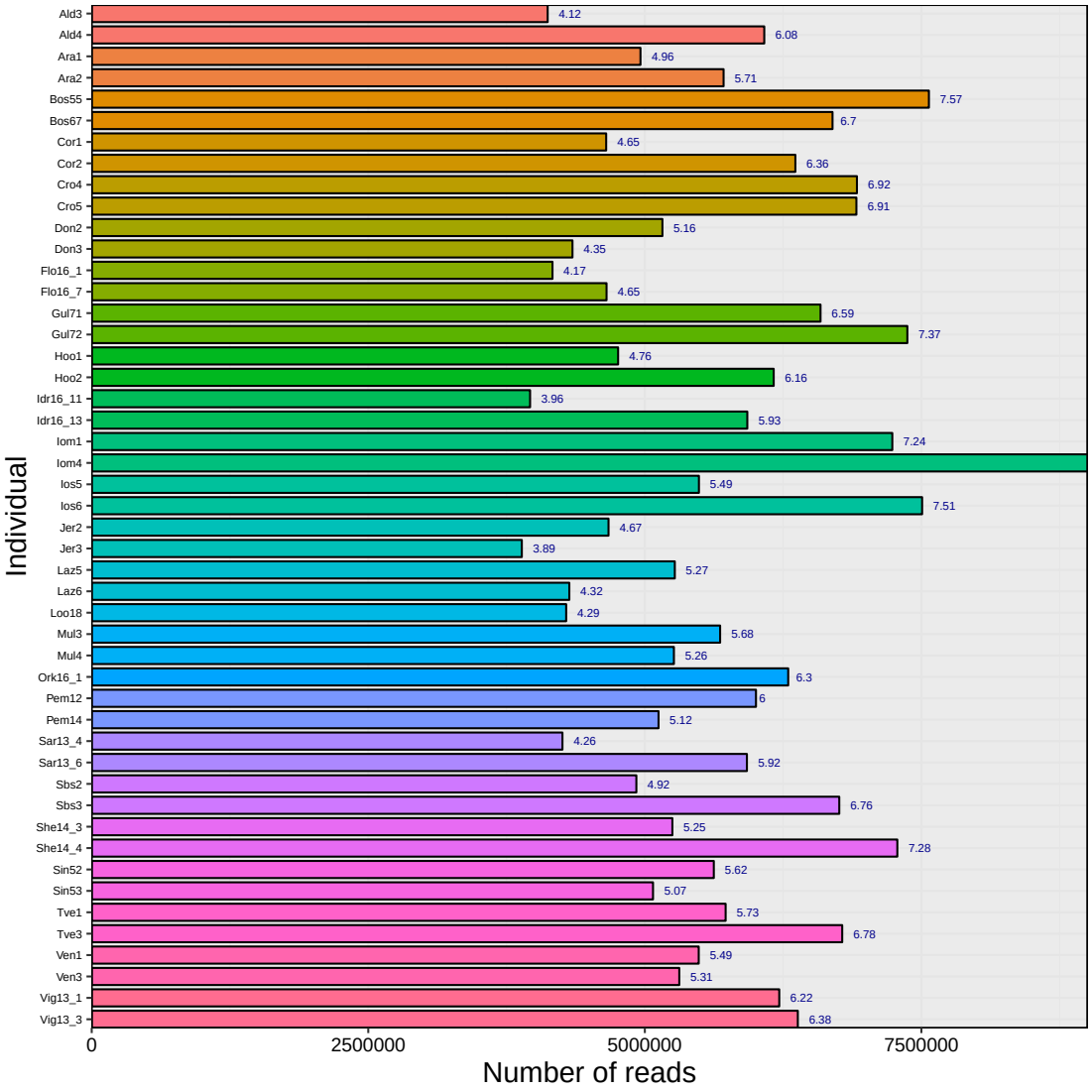


Figure 34: Total number of reads per individual.

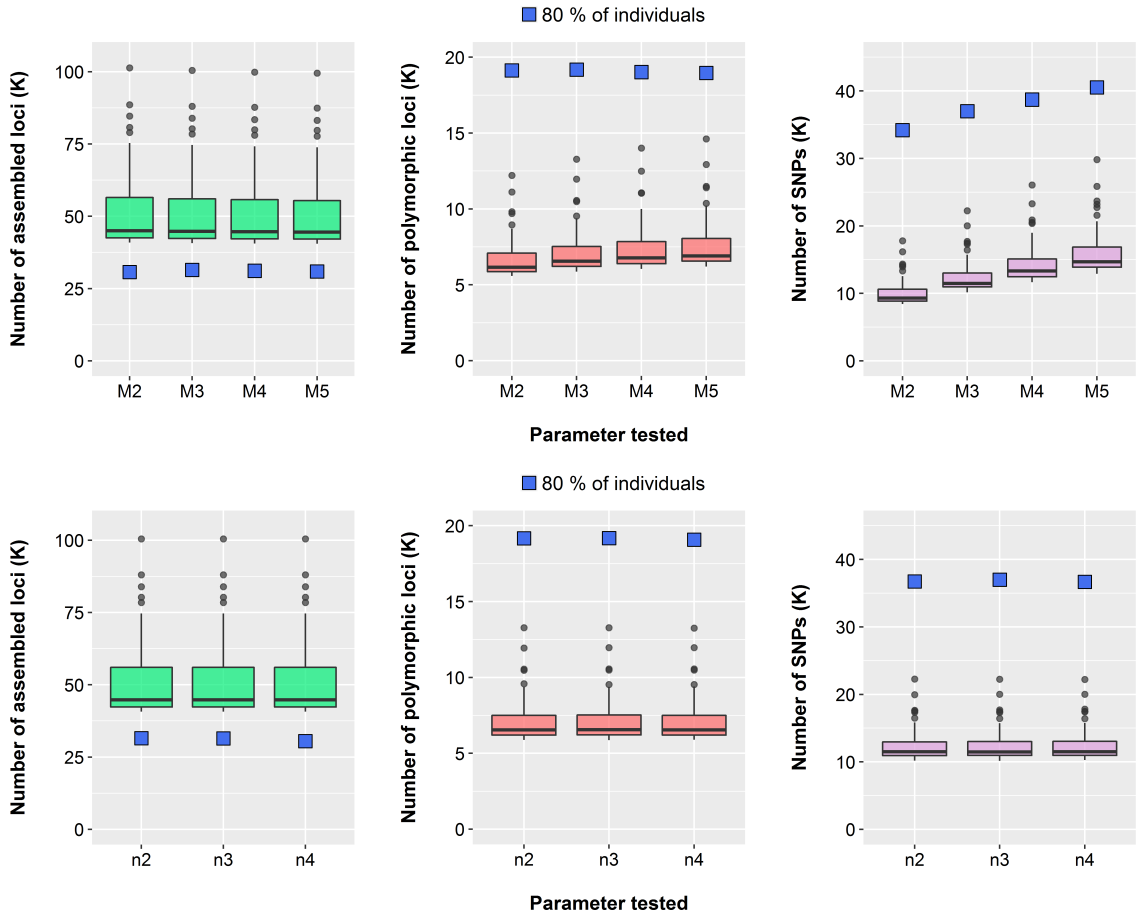


Figure 35: Exploring the optimum parameters for *de novo* assembly of RAD loci. The r80 represents the number of assembled loci, polymorphic loci or SNPs that are present in at least 80 % of the individuals.

5.3.3 SNP datasets

The SNP dataset exported from the `populations` program using all 55 samples consisted of 7,022 SNPs. Initial analysis using DAPC clustered these samples into three main groups: the Mediterranean (Italy samples), Skagerrak (Norway and Sweden samples), and the remaining samples from the northeast Atlantic (Fig. 36). Global F_{st} (Weir & Cockerham 1984) between these three groups was 0.018, calculated using the `diffCalc` function from diveRsity (Keenan et al. 2013).This was first evidence in this SNP dataset for genetic differentiation between the Atlantic and the Mediterranean in *H. gammarus*.

The SNP dataset exported from `populations` using only Atlantic samples (excluding Mediterranean and Skagerrak samples) consisted of 4,377 SNPs and 40 individuals organised into nine groups (Table 10); global F_{st} (Weir & Cockerham 1984) for this dataset was 0.002.

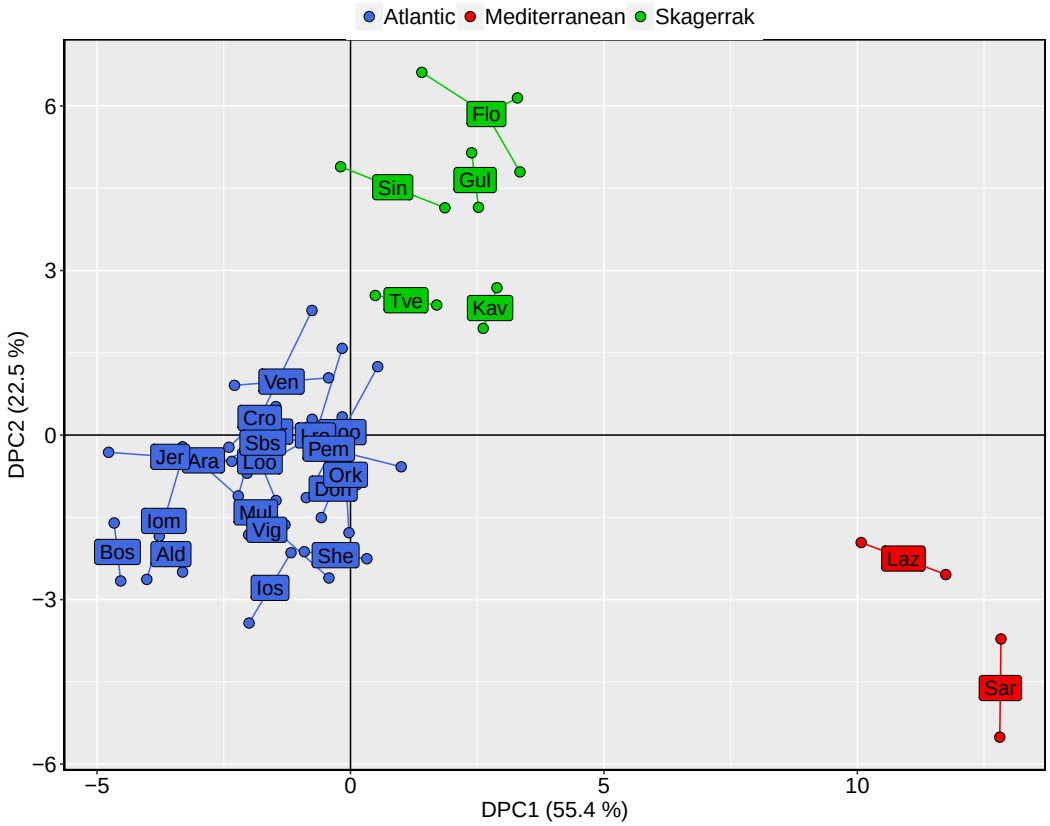


Figure 36: Discriminant analysis of principle components (DAPC) using 7,022 SNPs and 55 individuals.

5.3.4 SNP panel

The SNP panel was designed to contain SNPs with the highest G''_{st} from both datasets that were eligible for primer design and synthesis; this study was limited to 96 SNP markers due to assay development costs and the requirements of the Fluidigm EP1 system (i.e. dependency on 96-well plates). Using the dataset composed of all 55 individuals, and after filtering SNPs that were inadequate for primer design, 21 SNPs (out of 7,022 SNPs) were selected to capture differentiation between Atlantic, Mediterranean and Skagerrak lobsters (Fig. 37).

Using the dataset containing 40 Atlantic samples (excluding Mediterranean and Skagerrak samples), 75 SNPs (out of 4,377 SNPs) were selected with the aim of capturing differentiation between lobsters originating from geographically separate regions in the Atlantic.

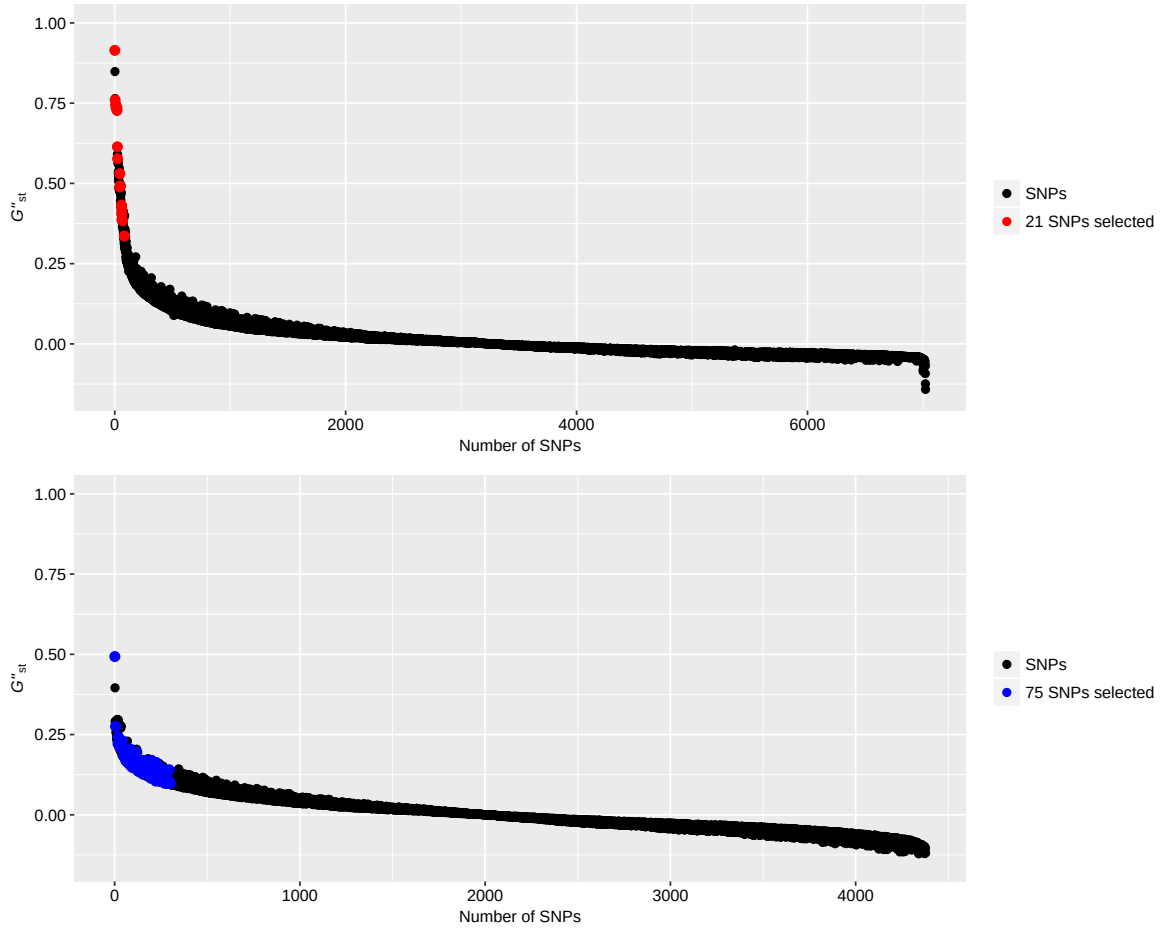


Figure 37: G''_{st} for each single nucleotide polymorphism (SNP) from two datasets: (top) SNPs ranked by G''_{st} when all samples were grouped by Atlantic, Mediterranean or Skagerrak; (bottom) SNPs ranked by G''_{st} using only Atlantic samples grouped by geographic region (Table 10). Red and blue points denote SNPs that were selected to compose the SNP panel.

5.4 Discussion

This study used RADseq to isolate SNP markers which capture genetic differentiation at different spatial scales across the range of *H. gammarus*. In particular, a panel of 96 SNP markers has been developed to capture differentiation between Atlantic and Mediterranean lobsters, and to capture hierarchical differentiation between Atlantic lobsters. The RADseq dataset in this study, although composed of extremely small sample sizes, provided a glimpse into the patterns of genetic structure in this species. A DAPC showed three distinct clusters: the Mediterranean, Skagerrak and the remaining northeast Atlantic samples (Fig. 36). However, although informative SNPs were selected to compose the SNP panel, only $\sim$0.30% of SNPs were retained from the original 7,022 SNPs, which may compromise power to detect these groups when further samples are genotyped. On the other hand, a very low number of individuals composed sample sites in the RADseq dataset, so an increase in the number of individuals per sampling site may increase the power to detect this differentiation. In any case, increasing the sample size for each site will facilitate more accurate and reliable calculations of allele frequencies.

Isolating markers from a small number of individuals using NGS can introduce effects of ascertainment bias (Helyar et al. 2011). Ascertainment bias results from the selection of loci from an unrepresentative sample of individuals that are then used to infer aspects of genetic variation and population structure across a broader part of the species' range that were not represented in the original SNP discovery step (Seeb et al. 2011). Acknowledging this potential bias is relevant for this study because several sites to be genotyped were sampled after the SNP discovery step. However, ascertainment bias was mitigated in this study by including samples in the SNP discovery step from across most of the current range of *H. gammarus*.

In studies of molecular ecology and population genetics, SNP genotyping is set to continuously increase in popularity as more markers are developed for model and non-model organisms (Seeb et al. 2011). Indeed, recently there has been an increase in the number of informative SNP panels developed, which are likely to have useful applications in forensic science (Martinsohn & Ogden 2009; Jacobs et al. 2018) and for the management of wild populations (Meek et al. 2016; Baetscher et al. 2017). When developing these SNP panels, considering the statistical power of the markers to detect genetic differentiation between spatially discrete populations is critical; this is mainly influenced by the number of SNPs composing the panel and the sample size of each population being studied, with an increase in the latter more likely to provide greater improvements in statistical

power (Morin et al. 2009). Simulations using hypothetical data have shown that there is high power (0.80) to detect a differentiation level of F_{st}=0.01 with 75 SNPs and with a sample size of 30 individuals per population (Morin et al. 2009). This suggests that with an increase in sample size per site, the SNP panel developed in this study for *H. gammarus* has high power to detect differentiation and resolve population structure. The next chapter uses this SNP panel to genotype additional lobsters from sites included in this RADseq study and from other sites sampled after the RADseq study.

Chapter 6: Exploring patterns of genetic structure, connectivity and assignment in the European lobster (*Homarus gammarus*) using SNP markers

---

Tom L. Jenkins and Jamie R. Stevens conceived and designed the study. Tom L. Jenkins coordinated and conducted sample collection, performed the SNP genotyping and analysed the data. Charlie D. Ellis and Alexandros Triantafyllidis assisted with sample collection. Tom L. Jenkins, Charlie D. Ellis and Mats B.O. Huserbraten designed the individual-based model (IBM) for European lobster larvae. Mats B.O. Huserbraten coded the IBM, performed simulations of larval dispersal and created the connectivity matrices.

6.0 Abstract

Delineating genetic structure and inferring connectivity in benthic marine invertebrates have been challenging due to typically weak genetic differentiation and the limited resolution offered by traditional genetic markers. In this study, the population genetic structure of the European lobster (*Homarus gammarus*), an economically important crustacean in the northeast Atlantic, was investigated across most of its current geographical range using a panel of genome-wide SNPs isolated using RADseq in a previous study. The main aims were to test: (i) whether fine-scale population structure exists across the *H. gammarus* range; (ii) whether patterns of connectivity are consistent with a stepping-stone model of dispersal; and (iii) whether this SNP panel can provide accurate individual assignment at various spatial scales, including ocean basin, region and sampling location of origin. After quality control and filtering, 1,223 lobsters from 36 sampling sites (38 including temporal samples) were genotyped at 86 biallelic SNP loci using a Fluidigm EP1 system. The results revealed strong genetic structure using all SNPs (global F_{st} = 0.06), partitioned between sites originating from the northeast Atlantic, the middle Mediterranean and the eastern Mediterranean (Aegean Sea). In the northeast Atlantic, there was a pronounced genetic cline starting from the most southerly sampling site (Vigo, northwest Spain) to the most northerly sampling site (Lysekil, western Sweden). Analysis indicated that isolation-by-distance is a key driver of this pattern; however, secondary contact after a period of isolation may also be responsible for this pattern. This was supported by the Bay of Biscay and northwest Spain harbouring the highest genetic diversity, suggesting that these two sites may have served as glacial refugia which preceded secondary introgression of northward dispersal after the Last Glacial Maximum. Individual assignment was 100 % accurate to basin of origin (i.e. Atlantic or Mediterranean); however, power was reduced for region of origin and significantly reduced for sampling location of origin. The findings of this study should be useful for lobster fisheries management, but also serve as another taxon to assess connectivity between MPAs in British waters.

6.1 Introduction

6.1.1 European lobster biology

The European lobster (*Homarus gammarus*) is a large decapod crustacean belonging to the family Nephropidae (clawed lobsters). Adults are usually territorial and found hiding in crevices within hard substrates composed of rock or

compressed mud in coastal areas from the low-tide mark to 150 m, but typically at depths not exceeding 50 metres. The current range of *H. gammarus* extends over most of the northeast Atlantic, from northern Norway to northern Morocco, and parts of the Mediterranean and the western Black Sea where they are found considerably more sparsely (Spanier et al. 2015). They are not found in the Baltic Sea, presumably because the larvae or adults cannot survive in the lower salinity and temperature conditions of the Baltic Sea (Jørstad et al. 2005), as is reported in some other marine invertebrates (e.g. Podbielski et al. 2016).

The genus *Homarus* is composed of *H. gammarus* and the closely related species, the American lobster (*Homarus americanus*), which occupies similar habitats in the western Atlantic, ranging from Cape Hatteras (North Carolina, USA) in the south to the Strait of Belle Isle (Labrador, Canada) in the north (Benestan et al. 2015). The next closest relative of these two clawed lobsters is the Norway lobster, *Nephrops norvegicus*, which was recently confirmed using data from complete mitogenomes (Shen et al. 2015). Homarid lobsters have two large specialised claws; one is blunt and designed for crushing, while the other claw is serrated and designed for slicing prey. The diet of adult lobsters, who mainly forage nocturnally, comprises of mostly benthic invertebrates such as crustaceans, molluscs, echinoderms and polychaetes, but can include algae, zooplankton and some fish. As with most crustaceans, the internal soft tissues are protected by a rigid exoskeleton, the front part of which forms the carapace and the hind portion comprises the abdomen and tail, which can be contracted quickly for rapid movements to escape danger.

Female *H. gammarus* reach sexual maturity in 5-7 years, at a carapace length (CL) of approximately 82-96 mm, but the size of sexual maturity can be geographically variable (Ellis et al. 2015a). Mating typically occurs in late summer and fertilised eggs are carried underneath the abdomen for 9-12 months while the embryos develop (Schmalenbach & Franke 2010); at this stage the female is said to be ‘berried’ (Fig. 38). Hatching typically occurs in spring-summer from late May to August and stage 1 larvae are discharged into the water column at night to begin their pelagic larval phase (Schmalenbach & Franke 2010).

The larvae are planktotrophic, feeding opportunistically on phytoplankton and zooplankton in the water column, and newly hatched larvae are thought to be positively phototaxic, which decreases as larvae age and develop (Schmalenbach & Buchholz 2010). Laboratory experiments have also shown that larvae are positively rheotactic, meaning they have some ability to swim against (horizontal or vertical) water currents (Schmalenbach & Buchholz 2010). The PLD is estimated to last 14-28 days and is dependent on sea temperatures encountered during

Figure 38: A berried European lobster.

transience (Schmalenbach & Franke 2010). The PLD is inversely proportional to temperature, meaning colder temperatures increase the PLD and vice versa; optimal larval survival in lobster larvae from Helgoland (southern North Sea) was found to be between 16°C and 22°C (Schmalenbach & Franke 2010). During this pelagic phase, larvae moult and develop into two intermediate stages, before free-swimming stage 4 post-larvae seek suitable benthic habitat to settle.

Contrary to *H. americanus*, the juvenile stage of *H. gammarus* is particularly cryptic and not well known. Juveniles of *H. americanus* have been found to favour cobble-boulder substrata (Wahle & Incze 1997); however, using a similar sampling method, no *H. gammarus* juveniles were found in this habitat across Norway, Ireland, Italy and the UK (Linnane et al. 2001). Common garden experiments suggest that *H. gammarus* juveniles may favour shelters that provide extensive tunnel systems (Jørstad et al. 2001; 2009), but exact nursery habitats in the wild still remain relatively unknown. Adult *H. gammarus* are thought to be relatively sedentary, with limited movement away from their home range (Skerritt et al. 2015). This is supported by several tagging studies, which demonstrate that adult movements are generally <3 km for periods up to one year (Oresland & Ulmestrand 2013; Skerritt et al. 2015), though some individuals have been recorded to travel up to 21 km away from their tagged origin (Huserbraten et al. 2013).

6.1.2 Fisheries

Global landings of the European lobster have been steadily increasing since the 1980s, with recent landings of 5,194 tonnes in 2014 (Fig. 39, FAO 2018). The American lobster fishery, in contrast, is significantly more productive, with recent global landings close to 160,000 tonnes in 2014. However, compared to the typical market price of the American lobster ($6.31-9.24 kg^{-1}), the European lobster has much higher value, fetching a price of £10.10-13.24 kg^{-1} at the time of writing (FIS 2018).

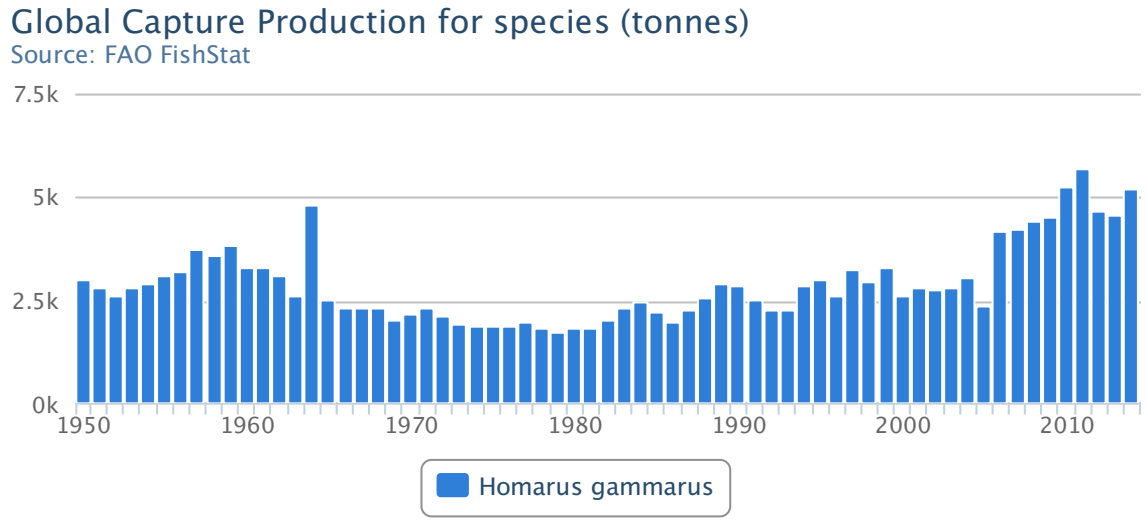


Figure 39: Global landings of European lobster (*Homarus gammarus*) from 1950-2014. Data from the Fisheries and Aquaculture Department of the Food and Agriculture Organization of the United Nations.

The high market value of *H. gammarus* makes it a prized seafood product for fishermen, so its fisheries are of great importance to the local and regional economies they support. However, current and historical over-exploitation has led to profound stock declines, with several regions (e.g. Scandinavia, Mediterranean, western Black Sea) experiencing stock collapses and from which recovery has been slow or stagnant (Kleiven et al. 2012). For example, in Norway, dramatic declines in landings from 1960-1980 indicated a severe collapse in the fisheries, and stocks have since remained at $<$10 % of their pre-1960 levels (Agnalt et al. 2007; 2009). Historically, *H. gammarus* was of particular interest to humans in parts of the Mediterranean and the western Black Sea; however, centuries worth of over-exploitation likely led to collapses and have possibly contributed to the current sporadic and scarce distribution and abundance of *H. gammarus* across these areas (Spanier et al. 2015).

The majority of landings of European lobster now originate from the coastal

fisheries of the UK, Channel Islands and the Republic of Ireland. In England, there are six Lobster Fishery Units (LFUs) (Fig. 40), which are reported to be based on the geographic distribution of the fisheries and what is known about hydrology and larval dispersal (CEFAS 2017). Each LFU spans one or more local Inshore Fisheries and Conservation Authorities (IFCAs) (Fig. 40), which manage coastal fisheries out to six nautical miles, whereas DEFRA and the Marine Management Organisation (MMO) are responsible for managing lobster fisheries beyond six nautical miles (CEFAS 2017). Local legislation can differ among the ten IFCAs; for example, in European waters there is a minimum landing size of 87 mm (CL), but several IFCAs (e.g. Devon Severn, Cornwall and Isles of Scilly) have introduced a larger minimum landing size of 90 mm. In Scotland, similar local regulations exist whereby, at the time of writing, most areas have a minimum landing size of 87mm, except for the Outer Hebrides, Orkney and Shetland where a minimum landing size of 90mm is enforced. This minimum landing size is linked to the age of maturity of females, in which most females (80-92 %, CEFAS 2017) this size are thought to be mature and have had the opportunity to spawn at least once. However, spatial and temporal variability and uncertainty in the size-maturity relationship in *H. gammarus* (Tully et al. 2001) has led to additional measures at some local and regional levels. For instance, it is illegal to land (i) lobsters that have been v-notched (a mark that remains for two-three moults), (ii) lobsters that have mutilated tail fans or (iii) lobsters that have been tagged, and in English waters it is also illegal to land berried females. The rationale behind this movement is to ensure a portion of the breeding female lobster population can be protected from fishing pressure. Moreover, in Scotland and the Republic of Ireland, there is a maximum size limit of 145mm (155mm in Orkney and Shetland) and 127mm, respectively. Female (and male) fecundity in *H. gammarus* has been found to be size-specific, with considerable variation between smaller spawners (CL 74mm, $\sim$4,000 eggs) and larger spawners (CL 151mm, $\sim$40,000 eggs) (Agnalt 2008); therefore, this legislation attempts to protect larger and more fecund females that contribute disproportionately to egg production.

Despite the legislation established for European lobster fishing, inaccurate landing reports from both commercial and recreational fishers can subvert these measures and undermine effective management of stocks (Kleiven et al. 2012). For example, in southeast Norway, Kleiven et al. (2012) found that only 24 % of lobster landed commercially were sold through the legal market and documented. The authors also found that recreational fishing, for which landings are unrecorded, accounts for 65 % of the total catch in the study area; however, whether this proportion is representative of other fishing areas across the northeast Atlantic is

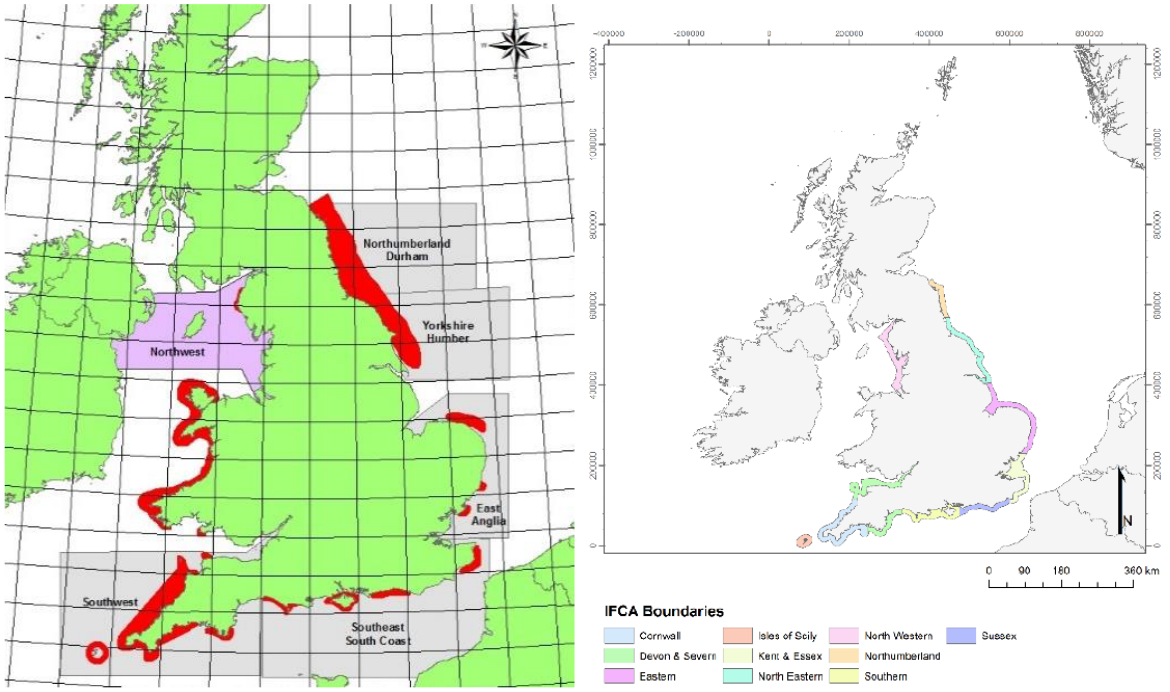


Figure 40: Lobster Fishery Units (left) and Inshore Fisheries and Conservation Authority boundaries (right) in England.

not known. In any case, this suggests that illegal, unreported and unregulated (IUU) fishing may seriously underestimate the actual landings of *H. gammarus*, and that many local and regional fisheries may have higher risks of over-exploitation than previously thought.

6.1.3 Hatchery stocking and aquaculture

Stock declines in several regions have led to the rearing of *H. gammarus* larvae in lobster hatcheries to produce juveniles which are released into the wild to supplement wild stocks (Bannister & Addison 1998; Ellis et al. 2015c). Such stocking has been implemented to either restore depleted or extirpated populations (restocking) or augment natural recruitment to maintain / increase yields (stock enhancement) (Bell et al. 2008). Initially, stock enhancement trial programs in the UK, France and Norway showed that some cultured juveniles can survive to adult sizes in the wild, which suggests that stocking may be a viable approach to augment natural stocks (Ellis et al. 2015c). However, uncertainty still surrounds whether stocking is economically viable, both in terms of the cost and labour required to produce cultured *H. gammarus* juveniles, and to monitor their performance in wild populations.

The most common method of producing juveniles has been to loan or purchase wild berried females from fishers or merchants and to hold the females in aquaria

until the eggs hatch (Ellis et al. 2015c). Larvae are then normally reared to at least developmental stage 5 (post-larval stages), and up to one year in some hatcheries, then released into natural habitats at an early benthic juvenile stage (Ellis et al. 2015c). This approach has released over 1.4 million cultured juveniles into European waters from 1983-2013, of which roughly 90 % of these releases were used for stock enhancement programs (UK, Ireland and France) and 10 % were used for restocking depleted populations (Norway, Germany and Italy) (Bannister & Addison 1998; Schmalenbach et al. 2011; Ellis et al. 2015c).

Recently, there have been technological advances in the aquaculture of European lobster which has seen juveniles grown in sea-based containers submerged in estuaries near to hatcheries (Beal et al. 2002; Benavente et al. 2010; Daniels et al. 2015). This mariculture approach is hoped to promote traits and behaviours that would develop in the natural environment and serve as an acclimation step before the juveniles are released (Ellis et al. 2015c). In addition, no supplemental feed is required as the juveniles consume natural prey in the environment. Such a project exists in southwest England, termed Lobster Grower (National Lobster Hatchery), whereby juveniles are reared in sea-based containers deposited in the estuary of the River Camel near Padstow. These approaches have massive potential to improve the ecological conditioning of cultured lobsters, which could increase their survival in the wild. Moreover, a future possibility of rearing juvenile lobsters to a marketable size could substantially reduce the pressure on wild populations. However, this approach is still in its infancy, and further cost and feasibility assessments are necessary to determine its economic viability (Daniels et al. 2015).

6.1.4 Previous genetic research

Genetic techniques offer the only approach for exploring the population structure and genetic diversity of *H. gammarus* stocks and for potentially determining the fate of hatchery-reared juveniles via parentage or population assignment (Gagnaire et al. 2015). Genetic research of paternity in both *H. gammarus* and *H. americanus* has yet to be applied to tracking juveniles bred from a hatchery; however, these studies have provided insights into mating patterns among males and females at different spatial scales (Gosselin et al. 2005; Ellis et al. 2015b; Sørdalen et al. 2018).

Population genetic structure has been investigated in *H. gammarus* using traditional molecular markers including random amplification of polymorphic DNA (RAPDs) (Ulrich et al. 2001), allozymes (Jørstad et al. 2005), mtDNA restriction fragment length polymorphisms (RFLPs) (Triantafyllidis et al. 2005) and

microsatellites (Huserbraten et al. 2013; Watson et al. 2016; Ellis et al. 2017). At regional and basin-wide scales, the general consensus using allozymes and mtDNA RFLPs suggests that lobsters from the Mediterranean, northern Norway and Oosterschelde (Netherlands) are genetically differentiated from each other and all other samples included in the studies (global theta=0.016, Jorstad et al. 2005; global G_{st}=0.078, Triantafyllidis et al. 2005). Ulrich et al. (2001) found similar patterns with fewer sampling sites and RAPD markers, and analysis with geographic distances suggested that this may be a product of IBD. Using 14 microsatellite markers, Ellis et al. (2017) were able to distinguish Skagerrak from all other sites sampled; however, these two sites from Skagerrak were genotyped by a different laboratory which suggests that artefacts from cross-calibration cannot be ruled out as a causal factor of the differentiation detected.

In comparison, at finer scales and using 12 microsatellites, high gene flow was found between sites within the Skagerrak region (Huserbraten et al. 2013) and between sites within the Irish Sea / Bristol Channel area (Watson et al. 2016). Moreover, Ellis et al. (2017) found high gene flow between all northeast Atlantic sites sampled except for Scandinavia. However, concerns over low sample sizes and geographical coverage, and limitations associated with the molecular markers used in these previous studies, question the power of these studies to adequately resolve the underlying population structure in this species. These potential limitations can be addressed by sampling more comprehensively across the current geographical range of *H. gammarus* and by investigating the genetic structure using informative SNP markers isolated from across the genome.

6.1.5 Study aims

The main aim of this study was to use the panel of SNPs developed by Jenkins et al. (2018b) to explore population genetic structure, connectivity and assignment in *H. gammarus* across our sampling sites. Specifically, this study asked three questions: (i) is there evidence of fine-scale population structure across the sampled range of *H. gammarus*?; (ii) are patterns of connectivity consistent with isolation-by-distance and a stepping-stone model of dispersal?; and (iii) can this panel of SNPs provide accurate assignment success at various spatial scales, including ocean basin of origin, region of origin and sampling location of origin. To conclude, future research objectives and implications for fisheries management and hatchery stocking are discussed.

6.2 Materials and methods

6.2.1 Sampling

Tissue samples of adult European lobsters were obtained in the same way as described in section 5.2.1. In total, 36 sites were sampled (38 including temporal samples), mainly during 2016 and 2017 (Table 11, Fig. 41). Because *H. gammarus* are extremely rare and precious in the Mediterranean, tissue and DNA samples collected in previous studies (Triantafyllidis et al. 2005; Ellis et al. 2017) were also included in this study (Table 11). Moreover, several Scandinavian samples collected in 2007 and 2009 were provided by Carl Andre (University of Gothenburg), which presented an opportunity to explore temporal genetic patterns in the Skagerrak area. This sampling strategy covered most of the current northeast Atlantic and Mediterranean distribution of *H. gammarus*, with only samples from northern Norway (northern limit), Portugal and Morocco (southern Atlantic), the western Mediterranean, and the western Black Sea (southern limit) absent, for which sourcing samples was attempted but was extremely difficult. All tissue samples were placed in 95-100 % ethanol and stored in a 4°C cold room for long-term preservation.

6.2.2 DNA extraction and SNP genotyping

Genomic DNA was extracted from all tissue types using a salting-out protocol (Appendix A4). The concentration and purity of all DNA extractions were quantified by spectrophotometry using a NanoDrop 1000. SNP genotyping was carried out on a Fluidigm EP1 system using the 96 SNPs isolated and developed by Jenkins et al. (2018b). Specific Target Amplification was carried out (as advised by Fluidigm) because it increases the copy numbers of the desired sequence containing the SNP, which can improve genotyping call rates and accuracy, particularly for heterozygous samples (Bhat et al. 2012). Assays and samples were run on a 96.96 Dynamic Array integrated fluidic circuit (IFC) and genotypes were called using the Fluidigm SNP Genotyping Analysis software (Fig. 42). A confidence threshold of 95 % was enforced and data were normalised using at least two negative controls per IFC run. The performance of the algorithm was checked after each analysis and obvious mistakes were amended; this included invalidating samples or assays that performed badly and correcting calls where the clustering algorithm was erroneous or ambiguous.

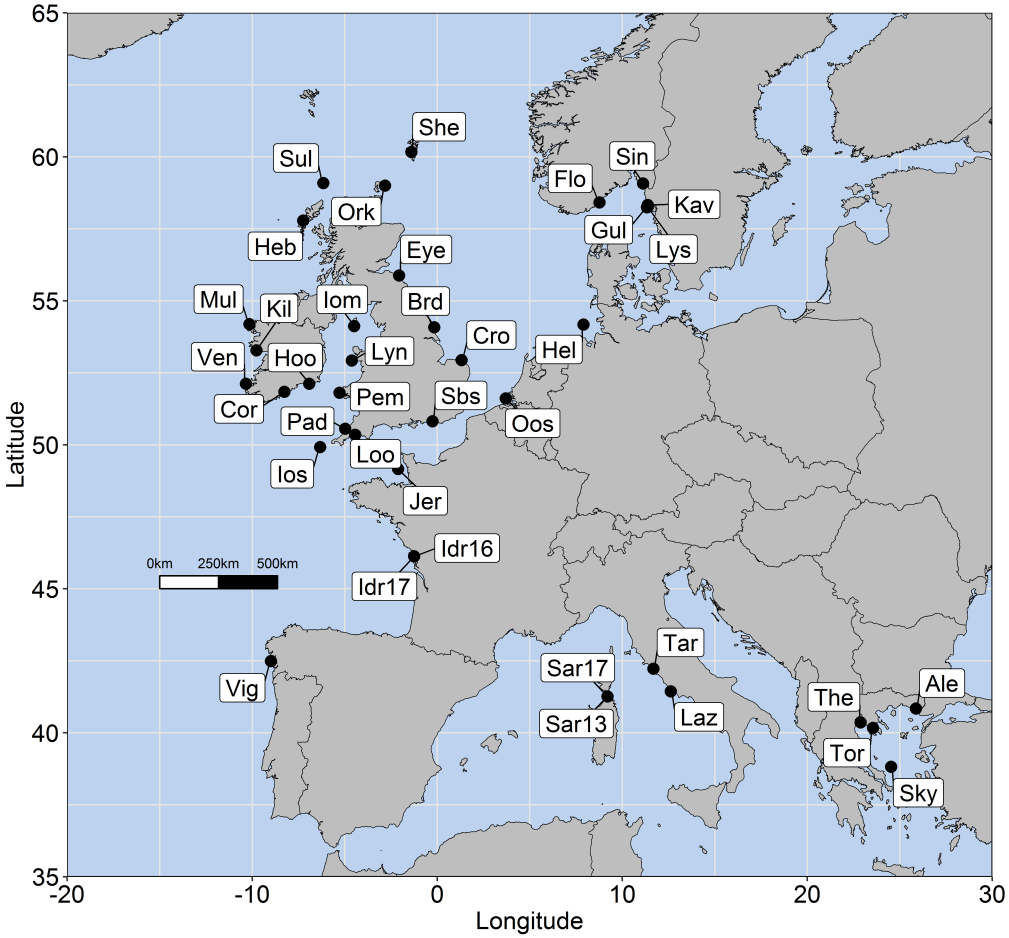


Figure 41: European lobster sampling sites.

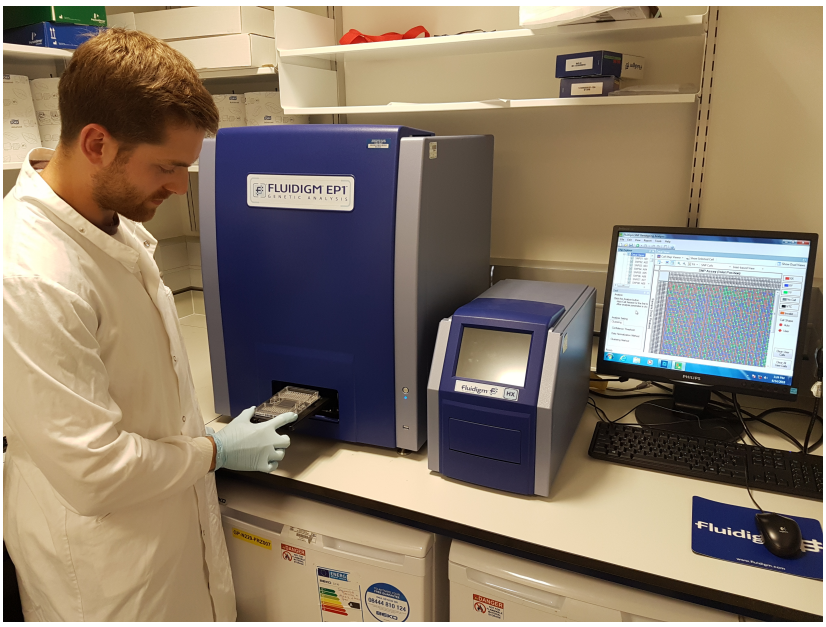


Figure 42: Fluidigm EP1 system.

Table 11: European lobster sampling information and genetic diversity statistics. F_{is} values significantly different from zero are highlighted in bold.

Country	Site	Code	N_1	N_2	Lat	Lon	Tissue type	Year	H_o	H_e	F_{is}
Britain	Bridington	Brd	36	36	54.07	-0.17	Pleopods	2017	0.36	0.35	-0.018
	[a]Cromer	Cro	36	35	52.94	1.31	Pleopods	2016	0.36	0.35	-0.027
	Eyemouth	Eye	36	27	55.88	-2.07	Pleopods	2017	0.37	0.36	-0.034
	Outer Hebrides	Heb	36	36	57.79	-7.25	Pleopods	2017	0.39	0.37	**-0.053**
	[b]Isle of Man	Iom	36	35	54.12	-4.50	Pleopods	2016	0.40	0.38	**-0.044**
	[a]Isles of Scilly	Ios	36	36	49.92	-6.33	Pleopods	2016	0.39	0.38	-0.015
	[a]Looe Harbour	Loo	36	36	50.35	-4.44	Pleopods	2016	0.39	0.37	**-0.066**
	[c]Llyn Peninsula	Lyn	36	34	52.93	-4.62	Pleopods	2017	0.41	0.38	**-0.068**
	Orkney	Ork	36	36	59.00	-2.83	Pleopods	2017	0.36	0.36	0.006
	Padstow	Pad	36	36	50.56	-4.98	Pleopods	2017	0.37	0.36	-0.023
	Pembrokeshire	Pem	36	36	51.81	-5.29	Pleopods	2016	0.38	0.37	-0.017
	Shetland	She	36	36	60.17	-1.40	Pleopods	2017	0.37	0.36	-0.025
	Shoreham-By-Sea	Sbs	36	36	50.82	-0.26	Pleopods	2016	0.37	0.36	-0.030
	Sula Sgeir	Sul	36	36	59.09	-6.16	Pleopods	2017	0.35	0.36	0.028
Channel Islands	Jersey	Jer	36	36	49.16	-2.12	Pleopods	2016	0.37	0.37	-0.002
France	Île de Ré, La Rochelle	Idr16	32	32	46.13	-1.25	V-notches	2016	0.38	0.38	-0.006
		Idr17	29	29	46.13	-1.25	V-notches	2017	0.40	0.39	-0.024
Germany	Helgoland	Hel	36	35	54.18	7.90	Pleopods	2017	0.33	0.33	-0.012
Greece	Alexandroupoli	Ale	35	28	40.84	25.87	DNA	1999-2001	0.33	0.34	0.040
	Skyros	Sky	37	37	38.82	24.53	DNA	1999-2001	0.35	0.34	-0.033
	Thermaikos Bay	The	37	36	40.36	22.88	DNA	1999-2001	0.35	0.34	-0.035
	Toronaios Bay	Tor	37	37	40.17	23.54	DNA	1999-2001	0.33	0.33	-0.001
Ireland	Cork	Cor	32	32	51.84	-8.26	Pleopods	2016	0.38	0.38	0.006
	[c]Hook Peninsula	Hoo	36	36	52.12	-6.92	V-notches	2016	0.39	0.37	-0.033
	[c]Kilkieran Bay	Kil	35	35	53.28	-9.77	Pleopods	2016	0.38	0.37	-0.031
	Mullet Peninsula	Mul	36	36	54.19	-10.15	V-notches	2016	0.37	0.38	0.016
	Ventry	Ven	36	36	52.12	-10.35	V-notches	2016	0.39	0.37	**-0.046**
Italy	Lazio	Laz	7	5	41.44	12.62	Antennae	2013	0.38	0.31	**-0.234**
	Tarquinia, Lazio	Tar	7	5	42.23	11.68	Antennae	2013	0.42	0.32	**-0.292**
	Sardinia	Sar13	7	7	41.26	9.20	Antennae	2013	0.32	0.29	-0.092
		Sar17	15	15	41.26	9.20	Pleopods	2017	0.34	0.34	-0.019
Netherlands	Oosterschelde	Oos	40	40	51.61	3.70	Pleopods	2017	0.31	0.32	0.010
Norway	Flodevigen	Flo	36	36	58.42	8.76	Pleopods	2016	0.34	0.33	-0.027
	Singlefjord	Sin	36	36	59.08	11.12	Pleopods	2009	0.34	0.33	**-0.041**
Spain	Vigo	Vig	36	36	42.49	-8.99	Pleopods	2017	0.40	0.39	-0.017
Sweden	Gullmarfjord	Gul	36	35	58.25	11.33	Pereiopods	2009	0.37	0.34	**-0.072**
	[d]Kavra	Kav	36	36	58.33	11.37	Pereiopods	2007	0.36	0.34	**-0.056**
	Lysekil	Lys	36	36	58.26	11.37	Pleopods	2017	0.32	0.33	0.014

N_1, number of individuals genotyped; N_2, number of individuals genotyped with missing data and duplicates removed; H_o, observed heterozygosity; H_e, expected heterozygosity; F_{is}, inbreeding coefficient.
[a]Marine Conservation Zone, [b]Ramsar site, [c]Special Area of Conservation, [d]Marine reserve.

6.2.3 Quality control and filtering

Individuals and SNP loci with more than 20 % missing genotypes were removed from the dataset using the `missingno` function from poppr v2.8.0 (Kamvar et al. 2014). Due to concerns over double-sampling, duplicated genotypes were identified using the `mlg` and `mlg.id` functions from poppr and were removed using custom R code. Deviations from Hardy-Weinberg equilibrium (HWE) were tested using the `hw.test` function from pegas v0.11 (Paradis 2010). The exact test based on Monte Carlo permutations of alleles was performed using 1,000 replicates. The false discovery rate (FDR), computed using the `p.adjust` function in R, was used to adjust for multiple comparisons; loci were considered to be out of HWE if they significantly ($p < 0.05$) deviated in more than 50 % of populations.

6.2.4 Genetic diversity

Observed heterozygosity (H_o), expected heterozygosity (H_e) and the inbreeding coefficient (F_{is}) were calculated using the `divBasic` function from diveRsity v1.9.90 (Keenan et al. 2013). Significance of (F_{is}) was assessed by calculating bias corrected 95 % confidence intervals (1,000 bootstrap replicates) and testing whether values were significantly different from zero.

6.2.5 Detecting outlier SNPs

Outlier SNPs potentially under divergent selection were identified using three differentiation-based methods. Firstly, BayeScan v2.1 (Foll & Gaggiotti 2008) was implemented using default parameters and a prior model (pr_odds) of 10,000, which sets the neutral model as being 10,000 times more likely than the model of selection to minimise the risk of false positives (Lotterhos and Whitlock 2014). Secondly, PCAdapt v4.0.3 (Luu et al. 2017) was run using three principal components (K=3). PCAdapt uses PCA to detect loci under selection and assumes that markers excessively related to population structure are candidates for local adaptation. Lastly, the infinite island model in Arlequin v3.5.2.2 (Excoffier & Lischer 2010) was run using 100,000 simulations and 1,000 demes. This method integrates heterozygosity and simulates a distribution of F_{st} for neutrally distributed markers. For all methods, a false discovery rate of 0.05 was used to identify outliers. Outlier tests were conducted on all filtered SNPs and, using the results of these tests, the SNPs were divided into three datasets: (i) all SNPs, (ii) putatively neutral SNPs, and (iii) outlier SNPs putatively under divergent selection (SNPs putatively under balancing selection were removed). A SNP was considered as

putatively under divergent selection if all three methods identified it as an outlier.

6.2.6 Population structure

Genetic differentiation between sampling sites was analysed by calculating pairwise values of F_{st} (Weir & Cockerham 1984) and D (Jost 2008) using the `diffCalc` function from diveRsity. Heatmaps of each statistic were visualised in R and significance was assessed using the same method previously described for F_{is}.

Population genetic structure was explored using two different approaches. Firstly, a DAPC was run using the `dapc` function from the R package adegenet v2.1.1 (Jombart & Ahmed 2011). Cross validation using the `xvalDapc` function from adegenet was used to choose the optimal number of PCs to retain. Secondly, STRUCTURE v2.3.4 (Pritchard et al. 2000), a Bayesian clustering algorithm, was run in parallel using the program StrAuto v1.0 (Chhatre & Emerson 2017). STRUCTURE was executed using the admixture model, with 10^5 MCMC repetitions and a burn-in of 10^5. The locprior option was selected, with sampling locations used as *a priori* information; all other parameters were set to default values. The maximum number of populations (K) assumed was 20 and ten independent replicates per K (1-20) were computed. To statistically assess different values of K, the mean value of L(K) (Pritchard et al. 2000) and the delta K (Evanno et al. 2005) statistics were examined in the R package pophelper v2.2.5.1 (Francis 2017). Replicates runs were aligned and merged with CLUMPP v1.1.2 (Jakobsson & Rosenberg 2007) using a wrapper script in pophelper and R was used to visualise the results.

6.2.7 Isolation-by-distance

As genetic connectivity between populations is driven by neutral processes, only putatively neutral SNPs were used for inferring dispersal and connectivity. For these analyses, temporal samples from both the Île de Ré (Idr16 and Idr17) and Sardinia (Sar13 and Sar17) were combined because of their genetic similarity; in addition, Laz and Tar were combined into one Lazio sample due to their spatial proximity and genetic similarity (see Results).

Two approaches were implemented that explore whether spatial distribution explains any of the observed genetic variation between sampling locations: (i) traditional Mantel tests and (ii) redundancy analysis (RDA). Mantel tests assume that the relationship between two dissimilarity matrices (D1:D2) is linear and that small D1 and large D1 values correspond to small D2 and large D2 values,

respectively. However, it has been suggested that these patterns rarely exist unless spatial correlation extends over the whole study area (Legendre et al. 2015). RDA is a combination of multiple linear regression and PCA that examines how much of the variation in a matrix of independent variables (i.e. spatial distribution, temperature, etc.) explains the variation in a matrix of dependant variables (i.e. allele frequencies). RDA assumes, like Mantel tests, that the expected relationship between the dependant and independent variables is linear. The main advantage of RDA over Mantel tests is that it operates on the raw data (i.e. the allele frequencies), as opposed to genetic distance matrices; this enables one to directly test how the spatial distribution of genetic variation within a species is influenced by effective dispersal (Meirmans 2015). In addition, RDA has been shown to be provide more power for assessing the influence of spatial correlation than Mantel tests (Legendre et al. 2015). It is also worth noting that the r^2 statistics of the Mantel test and from RDA are not comparable; in the Mantel test, r^2 measures the proportion of the dissimilarity variances in D1 that are explained by geographic distances, whereas RDA r^2 measures how much of the variance in the dependant variable (allele frequencies) is explained by geography (Legendre et al. 2015).

To create geographic distance matrices for the Mantel tests, least-cost distances (km) between sampling sites were calculated using the `lc.dist` function from the R package marmap v1.0 (Pante & Simon-Bouhet 2013). Matrices of pairwise F_{st} and D were calculated using the R package mmod v1.3.3 (Winter 2012) and negative values were converted to zeros. Mantel tests were performed using the `mantel.rtest` function from the R package ade4 v1.7.11 (Dray & Dufour 2007) and significance was assessed using 10,000 permutations. For RDA, geographic coordinates (lat and lon in decimal degrees) of sampling sites were transformed into Cartesian coordinates using the `geoXY` function from the R package SoDA v1.0.6 (Chambers 2013). Then, Euclidean distances were calculated from the Cartesian coordinates using the `dist` function and distance-based Moran's eigenvector maps (dbMEMs) were computed using the `dbmem` function from the R package adespatial v0.2.0 (Dray et al. 2018). The dbMEMs are a series of variables that summarise the spatial structure among the sampling sites, thereby representing a spectral decomposition of the spatial relationships between study sites (Borcard & Legendre 2002). RDA was performed on the population allele frequencies (dependant variables) and the dbMEMs (independent variables) using the `rda` function from the R package vegan v2.5.2 (Oksanen et al. 2018). Significance was assessed by analysis of variance (ANOVA) using the `anova.cca` function (10,000 permutations) and only significant ($p<0.05$) dbMEMs were included in subsequent analyses.

6.2.8 Simulations of larval dispersal

Oceanographic drift modelling was implemented to estimate larval dispersal and predict connectivity between study sites for a comparison with the observed patterns of genetic connectivity. The hydrodynamic model used is described in Lien et al. (2014) and the particle-tracking algorithms applied are detailed in Vikebø et al. (2010). The ocean current model had a horizontal resolution of 4 km and simulations were performed from 2007-2016, except for 2012-2014 for which data were unavailable; however, patterns of connectivity were consistent across the years, so the results were deemed representative (Mats Huserbraten, *pers. comm.*). Due to the limited spatial extent of the ocean model, only study sites from the English Channel to the North Sea and Skagerrak were included in the simulations; as it represents a partially closed system, Oosterschelde was also outside the scope of the ocean model so the closest feasible point in the North Sea was used instead. Furthermore, Gullmarfjord, Kavra and Lysekil (western Sweden) were located in the same grid-cell and were therefore merged in the simulations.

An individual-based model (IBM) was designed for European lobster larvae that incorporated aspects of their known life history. Particle (larvae) release followed a normal distribution with peak release in mid-June; approximately 6,000 particles were released from each site across all years, which translated to a total of 1 million particles released over 200 days. Particles drifted at a fixed depth between 1-20 m for 12-28 days depending on the median temperature encountered during the drift trajectory (Schmalenbach & Franke 2010). The PLD decreased with increasing temperature and increased with decreasing temperature; however, there is a critical value in which temperature affects survival in hatchery-reared lobsters, resulting in mortality (Schmalenbach and Franke 2010). Therefore, when larvae encountered median temperatures of $<$14°C and $>$22°C (i.e. larvae experienced temperatures less than $<$14°C or $>$22°C for more than half of their drift time), larvae were considered dead because the temperature was assumed to be outside of the thermal niche required for development.

Dispersal trajectories for all sites and years were plotted with and without the use of the IBM to compare dispersal patterns including and excluding biological parameters (i.e. temperature-dependency and mortality). Connectivity between sites was assessed by creating a connectivity matrix, whereby one unit in the matrix represented one day spent by a source particle within a 40 km radius from a sink.

6.2.9 Individual assignment

The accuracy of assigning individuals back to their basin of origin (i.e. Atlantic Ocean or Mediterranean Sea) and to their sampling location/region of origin was assessed using the R package assignPOP v1.1.4 (Chen et al. 2018). assignPOP uses a cross validation procedure followed by PCA to evaluate assignment accuracy and membership probabilities. Firstly, the dataset is partitioned into training (baseline) and test (holdout) datasets using a resampling cross validation procedure, with the user specifying the number or proportion of individuals from each source 'population' (i.e. Atlantic or Mediterranean in the basin analysis) to be used in the training dataset. This approach of creating randomly selected, independent training and test datasets avoids introducing high-grading bias (Anderson 2010). Secondly, the features of the training datasets (i.e. the genotypes) are reduced in dimensionality using PCA, the output of which are used to build predictive models from user-chosen classification machine-learning functions (Chen et al. 2018). Finally, these models are then used to estimate membership probabilities of test individuals and assign individuals to a source population, while also evaluating the baseline data and conducting assignment tests on individuals for which the origin is unknown (Chen et al. 2018).

All filtered SNPs were used in the assignment tests because the inclusion of both neutral and outlier loci can increase the power of assignment tests (Gagnaire et al. 2015). For assigning individuals to their basin of origin, before dividing the dataset into baseline and test datasets, two individuals per sampling location (76 individuals in total) were randomly selected in R to compose a file representing 'unknown' individuals, whereby the basin of origin was considered to be unknown. Due to the potential bias of unequal sample size in assignment studies (Wang et al. 2016), 250 individuals from the Atlantic basin were randomly selected to compose this source population (with 154 individuals composing the Mediterranean basin).

A Monte-Carlo cross validation procedure was used to group individuals into baseline and test datasets using the function `assign.MC` from assignPOP. Resampling was repeated 30 times for each combination of training individuals and loci. The proportion of individuals from each source population randomly allocated to the baseline dataset was set to 0.5, 0.7 and 0.9. Lastly, the support vector machine (svm) and the linear discriminant analysis (lda) classification functions were used to build predictive models. After building predictive models based on the baseline dataset, the origin of the unknown individuals were assessed, further testing the accuracy of the assignment.

6.3 Results

6.3.1 SNP genotyping and quality control

Five SNP assays (H_gam_25580, H_gam_32362, H_gam_41521, H_gam_53889, H_gam_65376) did not work consistently on the Fluidigm EP1 system, possibly due to inadequate assay design, poor STA amplification, or ascertainment bias. One locus (H_gam_22365) contained 28.3 % missing data and was therefore removed from the dataset. In addition, eight individuals (Ale04, Ale06, Ale08, Ale13, Ale15, Ale16, Ale19 and The24) were removed because of missing data ranging from 42.9-54.9 %, which was likely due to very poor DNA quality, evidenced by gel electrophoresis, because repeats also produced similar levels of missing data.

In total, 1,223 unique multi-locus genotypes from 1,242 individuals were obtained (Table 11). Although duplicates were apparent between some Laz and Tar samples (western Italy), most duplicates were mainly individuals from the same sampling site (Table 12). These results could have arisen from double-sampling during sample collection, contamination of DNA samples, or the inclusion of closely related siblings. In any case, because the exact cause could not be determined, only one individual from each duplicate was retained. One locus (H_gam_21197) significantly deviated from HWE and was removed from the dataset. In addition, three loci (H_gam_8953, H_gam_21880, H_gam_22323) exhibited an unexpectedly high proportion of observed heterozygosity (0.72, 0.61 and 0.67, respectively); these loci were removed because they could be paralogous loci as true variants are often considered to have a frequency of 0.50 heterozygous genotypes (Dufresne et al. 2014). The final filtered dataset contained 1,223 individual lobsters from 36 sampling sites (38 including temporal samples) and 86 biallelic SNP loci.

Out of the 96 SNPs selected in Chapter 5, all ten SNP loci that were omitted after quality control in this study originated from the 75 SNPs selected from 4,377 SNP dataset (Fig. 37). This meant that only 65 of the 75 SNPs were used in this study, while all 21 SNPs selected from the 7,022 dataset (Fig. 37) were retained as they met the quality control thresholds.

6.3.2 Genetic diversity

The H_o and H_e ranged from 0.31-0.42 (Oos-Tar) and 0.29-0.39 (Sar13-Idr17 and Vig), respectively (Table 11). Overall, lower diversity measures (H_e) were found in Oosterschelde, Helgoland, Scandinavia and the Mediterranean samples, while the

Table 12: European lobster individual IDs with identical genotypes.

Sampling site	Individual 1	Individual 2	Individual 3
Cromer	Cro08	Cro15	–
Eyemouth	Eye01	Eye17	–
	Eye02	Eye04	–
	Eye05	Eye06	Eye23
	Eye07	Eye24	–
	Eye14	Eye31	–
	Eye15	Eye16	–
	Eye20	Eye36	–
	Eye25	Eye29	–
Isle of Man	Iom02	Iom22	–
Gullmarfjord	Gul86	Gul101	–
Helgoland	Hel07	Hel09	–
Lazio / Tarquinia	Laz01	Tar01	–
	Laz02	Tar02	–
	Laz03	Tar03	–
	Laz04	Tar04	–
Llyn Peninsula	Lyn04	Lyn15	Lyn34

highest diversity measures were found in northwest Spain (Vig), the Bay of Biscay (Idr), and a few western and southerly sites from Britain and Ireland. The inbreeding coefficient (F_{is}) ranged from -0.292 (Tar) to 0.040 (Ale), and ten F_{is} measures were significantly different from zero. Laz, Sar13 and Tar had the lowest significant F_{is} measures (-0.234, -0.292 and -0.092, respectively), which were also the sites with very low sample sizes (Table 11). The lowest significant measures excluding these sites were -0.072 (Gul) and -0.068 (Lyn).

6.3.3 Outlier SNP detection

Bayescan and Arlequin detected 17 SNPs putatively under divergent selection, while PCAdapt detected 22 SNPs putatively under divergent selection. All three methods identified the same 15 SNPs as outliers potentially under divergent selection; the remaining SNPs were considered neutral. In addition, outlier tests were also performed on the original RADseq dataset composed of 7,022 SNPs (Chapter 5); these analyses identified 59-124 outlier SNPs depending on the outlier test used, of which 13 out of the 15 outliers identified in the SNP dataset from this study were also identified in at one least or more outlier tests in the RADseq dataset.

6.3.4 Population structure

Global values of F_{st} and D were 0.060 and 0.014, respectively, and both pairwise differentiation statistics showed comparable pairwise patterns between sampling sites (Fig. 43). Values of F_{st} ranged from zero to 0.314 (Oos-Sar13) and from zero to 0.041 (Oos-Sar13) for D. The highest values for both statistics were between the Mediterranean sites and the Atlantic sites, of which many of these values were significantly different from zero. The lowest values tended to be between sites originating from Britain, Ireland, France and the Channel Islands. However, sites situated spatially close together in other regions also had low pairwise comparisons (i.e. Laz, Sar and Tar in the mid-Mediterranean; Ale, Sky, The and Tor in the eastern Mediterranean; and Flo, Gul, Kav, Lys and Sin in Skagerrak).

The DAPC using all 86 SNPs showed distinct separation of lobsters from the Atlantic and the Mediterranean (Fig. 44). There was also evidence for structure within the Mediterranean, partitioned between the mid-Mediterranean (Sar, Laz and Tar) and the eastern Mediterranean (Aegean Sea – Ale, Sky, The and Tor). Within the Atlantic cluster, there was a pronounced genetic cline starting from the most southerly site in the northeast Atlantic (Vig) to the most northerly sites in Skagerrak (Kav, Flo, Gul, Lys and Sin). In total, the first and second axes explained 69.8 % of the variation in the dataset. The 15 outlier SNPs showed very similar patterns to those described using all SNPs (Fig. 44); however, the first and second axes explained considerably more of the variation in the dataset (91.3 %). In addition, compared to the DAPC using all SNPs, differentiation between the middle and eastern Mediterranean was weaker.

The 71 neutral SNPs, in contrast, showed the same Atlantic-Mediterranean divide but the separation was weaker, particularly between mid-Mediterranean sites and Atlantic sites (Fig. 45). There was also stronger separation between the middle and eastern Mediterranean; however, overall the first three axes explained a lower amount of variation in the dataset. Furthermore, in comparison to the other SNP datasets, the Atlantic cluster showed no obvious clinal pattern using the neutral SNPs, but weak separation of Oosterschelde lobsters from the main Atlantic cluster was apparent after exploring and visualising the first three axes (Fig. 45).

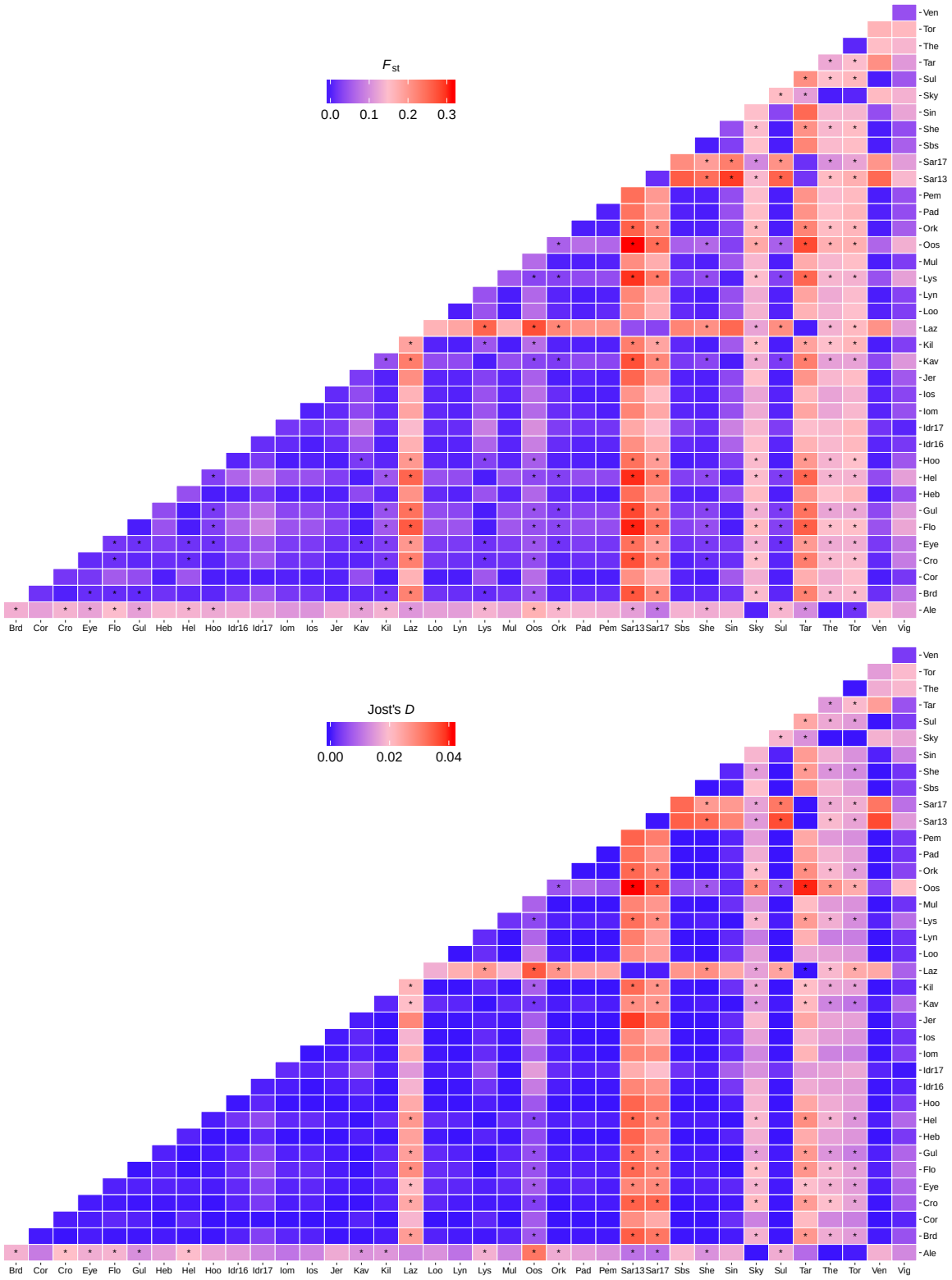


Figure 43: European lobster heatmaps of F_{st} (top) and D (bottom). Asterisks represent values significantly different from zero.

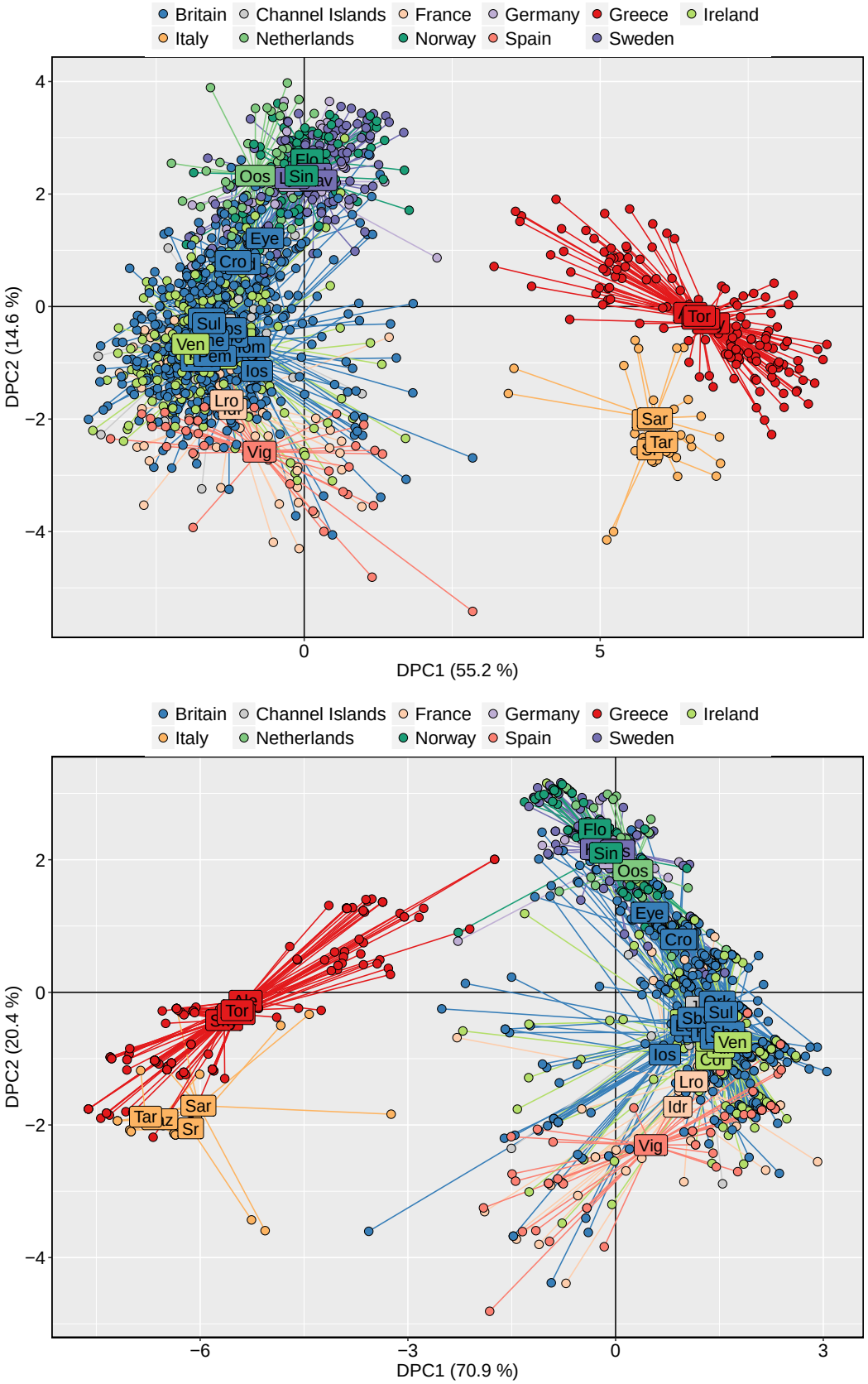


Figure 44: Discriminant analysis of principal components using all 86 SNPs (top) and 15 outlier SNPs (bottom). For each DAPC, points represent individuals and colours denote the sampling region of origin.

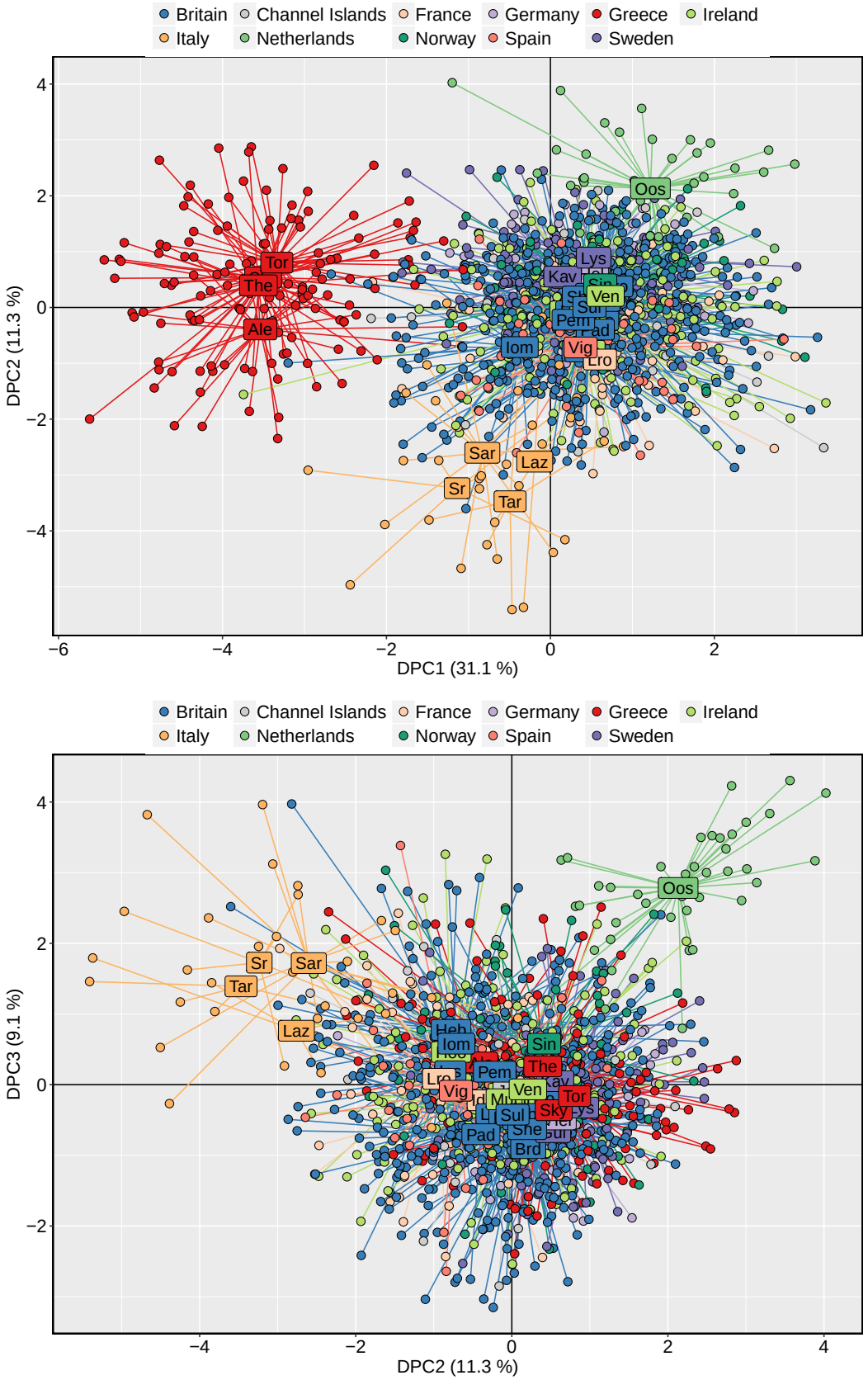


Figure 45: Discriminant analysis of principal components using 71 neutral SNPs: axis 1 and 2 (top) and axis 2 and 3 (bottom) are shown. For each DAPC, points represent individuals and colours denote the sampling region of origin.

STRUCTURE analysis generally supported the DAPC results for all SNP datasets (all SNPs, neutral SNPs and outlier SNPs); however, a genetic cline was evident from the eastern Mediterranean to Skagerrak in all SNP datasets (Fig. 46). After exploring different levels of *K* and examining the statistics (Appendix A7), the most likely *K* for all SNP datasets was *K*=3, which was generally consistent with the clusters found from the DAPCs. The membership proportions were estimated for each individual (Fig. 47, 48) and the mean was calculated for each sampling site to generate an average membership proportion to each *K* cluster which was visualised as pie charts on a map (Fig. 46). In the neutral SNP dataset, Oosterschelde was not apparent in *K*3; however, analysis of *K*5 showed that one cluster was predominantly found in Oosterschelde, which was also consistent with the results from the neutral DAPC.

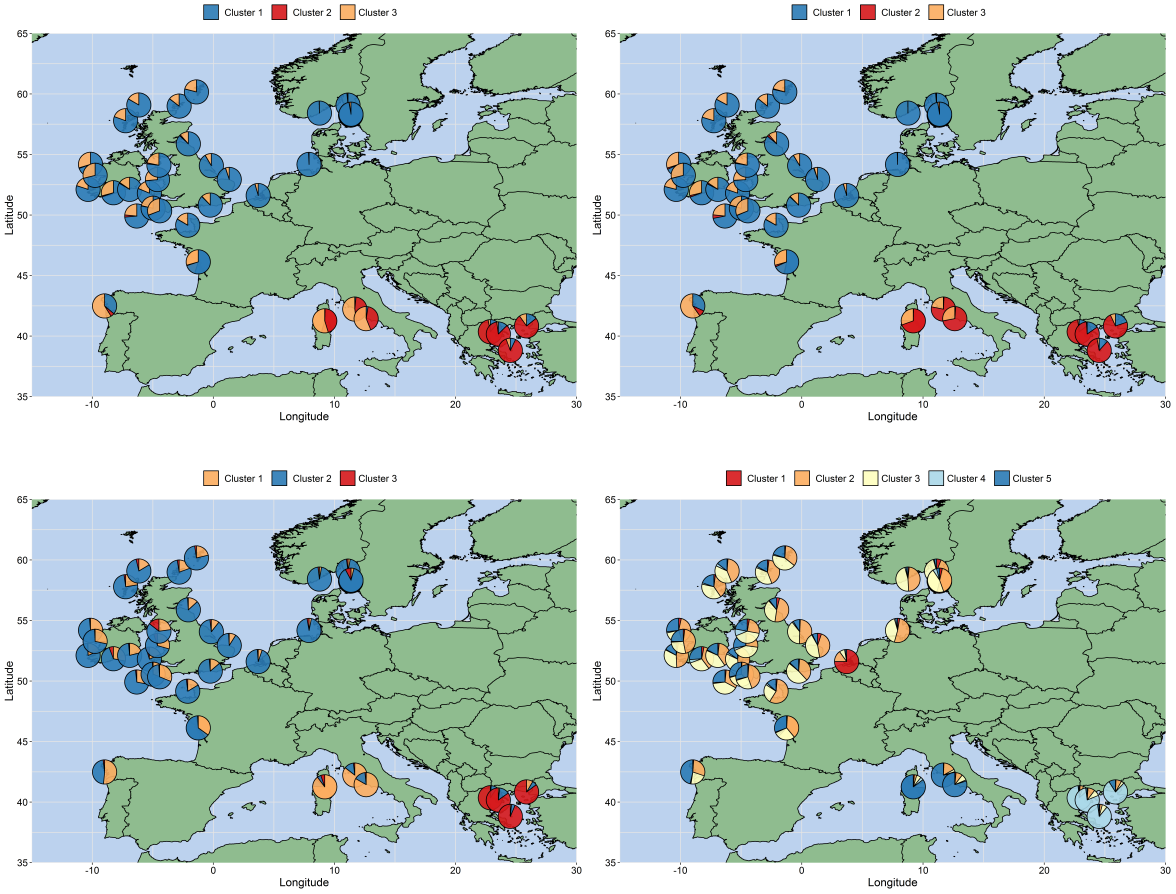


Figure 46: STRUCTURE results per individual using all 86 SNPs (top-left, *K*3), 15 outlier SNPs (top-right, *K*3), 71 neutral SNPs (bottom-left, *K*3; bottom-right, *K*5). The mean was calculated for each sampling site to generate an average membership proportion to each *K* cluster which was visualised as pie charts on a map.

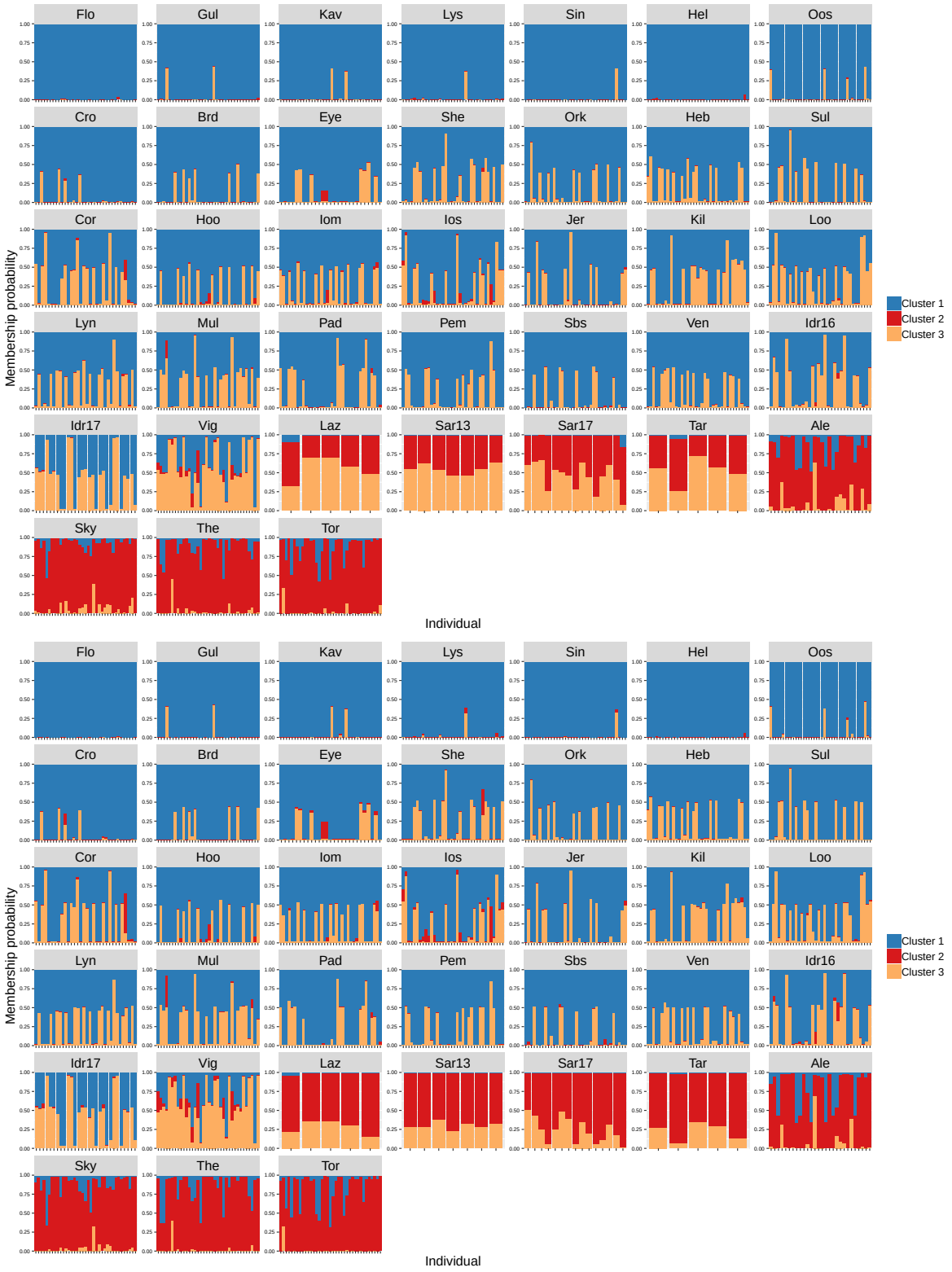


Figure 47: STRUCTURE results per individual using all 86 SNPs (top, *K*3) and 15 outlier SNPs (bottom, *K*3). Each bar represents an individual and colours denote membership proportions to each cluster.

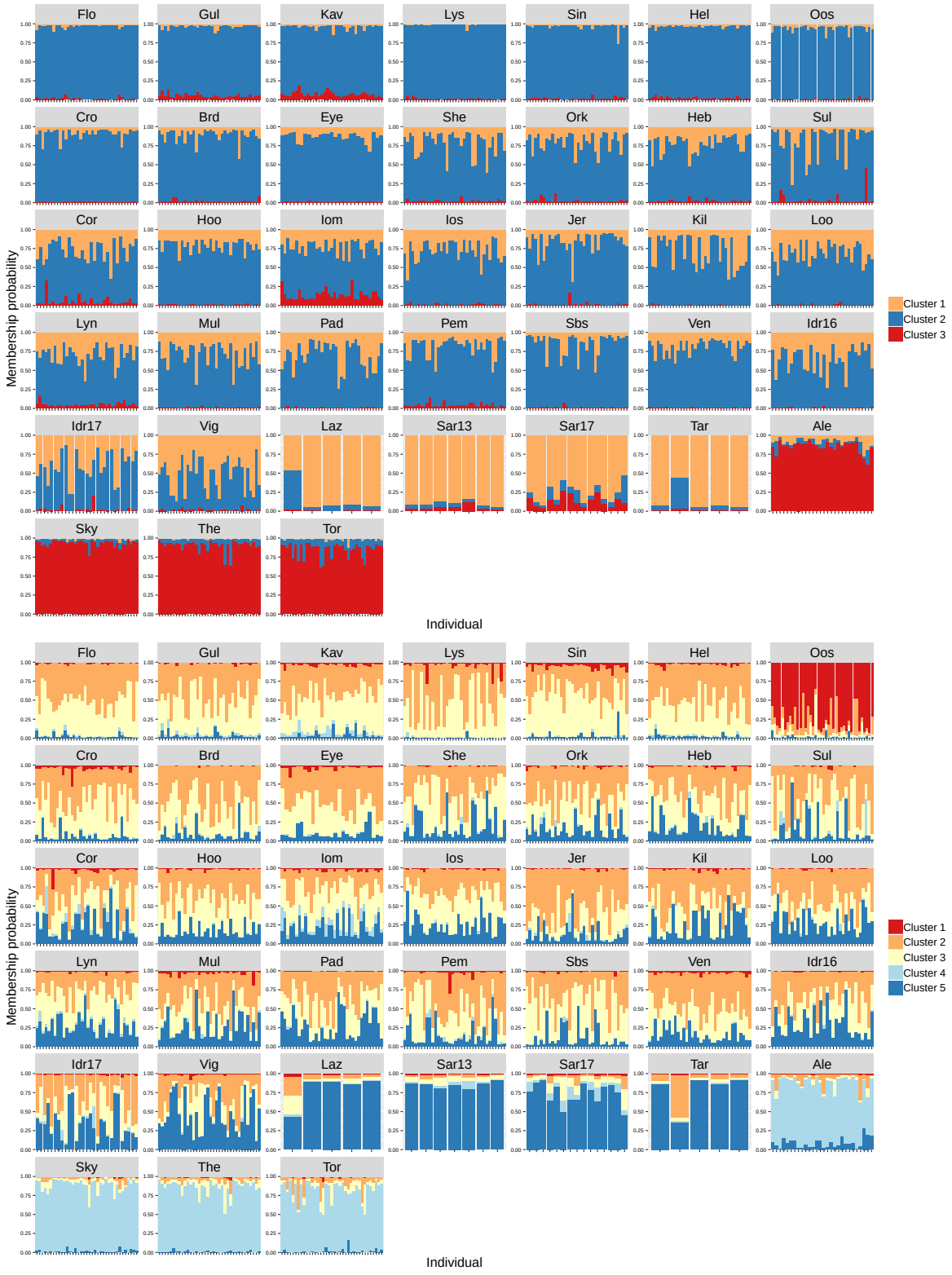


Figure 48: STRUCTURE results per individual using 71 neutral SNPs (top, *K*3; bottom, *K*5). Each bar represents an individual and colours denote membership proportions to each cluster.

To further visualise differentiation between the three groups found by DAPC and STRUCTURE, the population allele frequency for one allele was visualised for each SNP identified as an outlier in this study (Fig. 49). Most of these SNPs showed large differences in frequency between the Atlantic, the middle Mediterranean and/or the eastern Mediterranean. For instance, SNPs *22740*, *33066*, *51507*, *53052*, *53263*, *65064* and *65576* showed noticeably different frequencies between sites from the middle Mediterranean compared to sites from both the Atlantic and the eastern Mediterranean. In addition, SNPs *42395* and *53935* were completely fixed for the T allele in the eastern Mediterranean, while both SNPs were fixed in most, but not all, of the sampling sites in the middle Mediterranean. In the Atlantic, SNP *58053* was fixed for the A allele in several sites (typically sites from Scandinavia), while sites from both the middle and the eastern Mediterranean had comparably very low frequencies of the A allele.

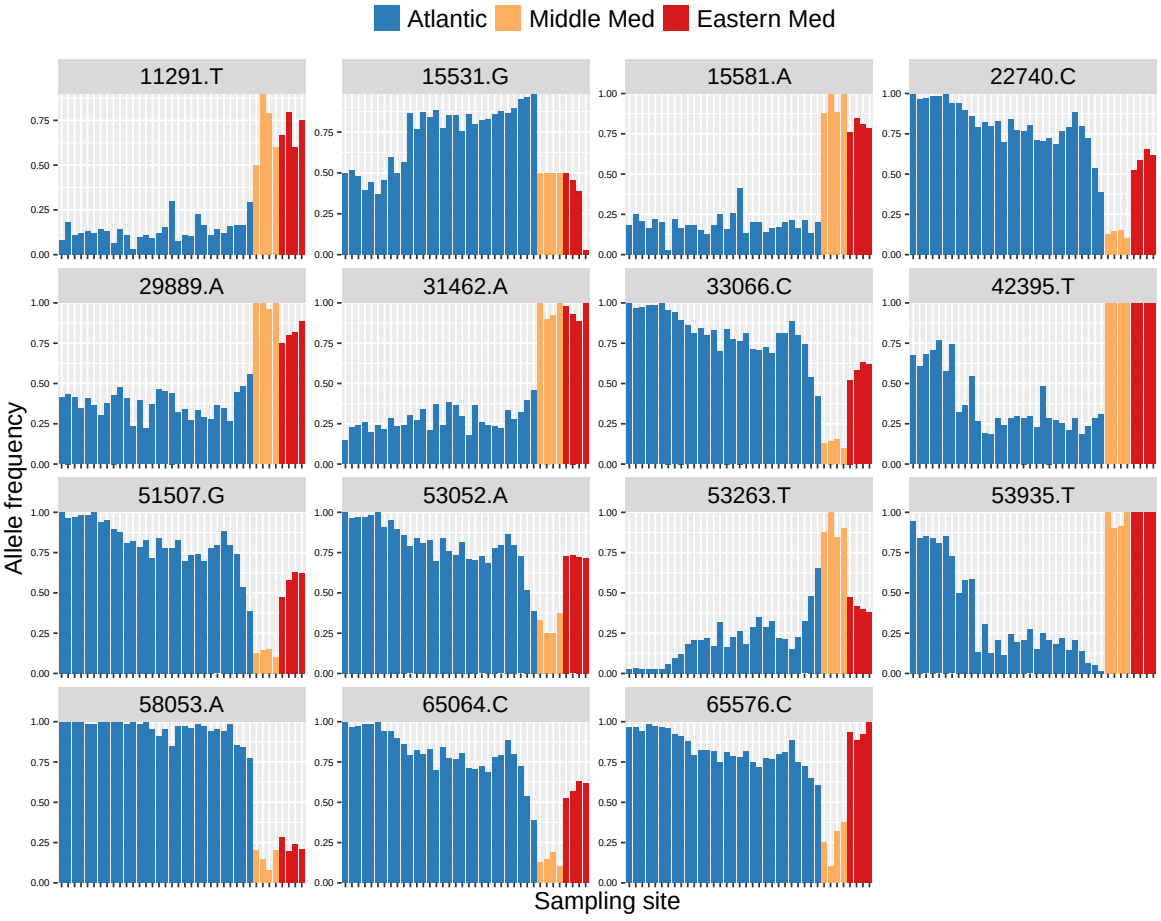


Figure 49: Population allele frequency of one allele for each outlier SNP identified in this study. For each SNP, the sampling sites (x axis) are arranged in the same order as the STRUCTURE results (Fig. 47, 48). Colours denote the sampling site region of origin: the Atlantic (blue), the middle Mediterranean (orange) and the eastern Mediterranean (red).

6.3.5 Isolation-by-distance

Initial analysis with Mantel tests showed concordant patterns using F_{st} and D; thus, only results for F_{st} are described. Using all sites, there was a strong positive correlation between geographic distance and F_{st} (Fig. 50a; r^2=0.81, p<0.001). When the Mediterranean samples were removed, this correlation was much weaker, but still significant (Fig. 50b; r^2=0.17, p=0.041). However, the removal of Oosterschelde lobsters vastly increased the strength and significance of the correlation (Fig. 50c, r^2=0.49, p<0.001). Analysis with only the Mediterranean samples also produced a strong correlation, but this was not significant (Fig. 50d; r^2=0.88, p=0.062).

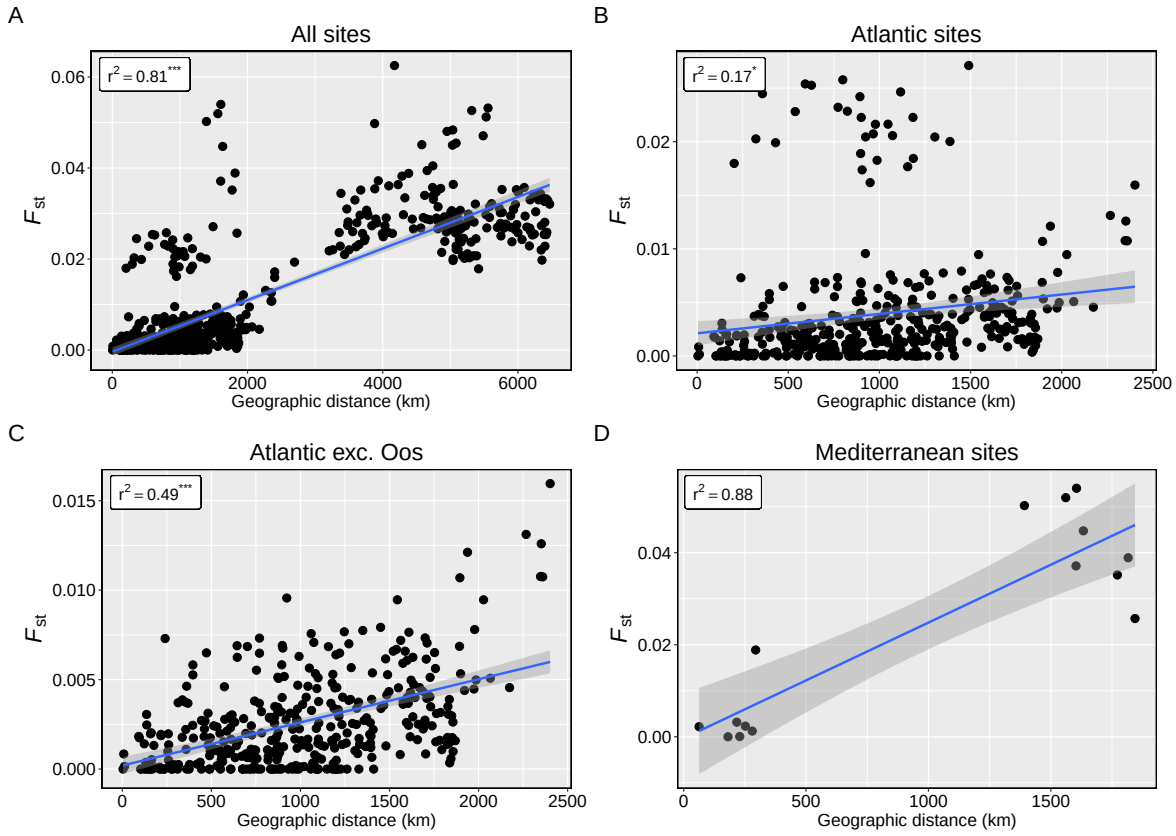


Figure 50: European lobster isolation-by-distance analyses: pairwise comparisons of geographic distances (km) and F_{st} between sampling sites were plotted using all sites (top-left), Atlantic sites only (top-right), Atlantic sites only and excluding Oosterschelde (bottom-left), and Mediterranean sites only (bottom-right). A linear regression line (blue line) and the standard error (dark grey shaded area) was added to each plot. Asterisks denote significance levels: *<0.05, **<0.01, ***<0.001.

The RDA was globally significant (R^2=0.43, p<0.001) and the first two axes of the RDA accounted for 52.1 % and 29.1 %, respectively. Of the eight dbMEMs constructed from the Euclidean distances, only vectors 1-2 were found to be significant for explaining variation in the allele frequencies. RDA was re-run with

these two dbMEMs and the results indicated that these two spatial variables explained 28.0 % of the variation in the allele frequencies (Fig. 51).

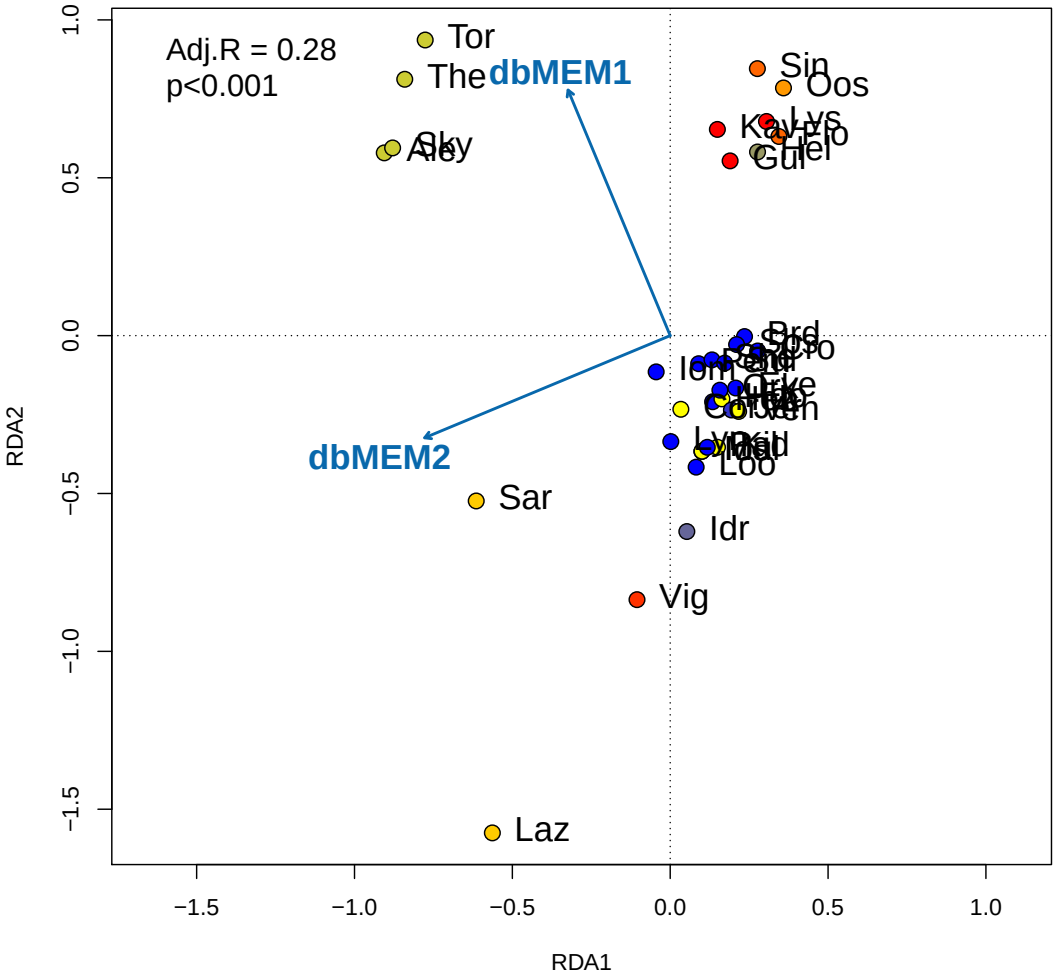


Figure 51: European lobster redundancy analysis. Each point represents a sampling site. Blue lines show the distance-based Moran's eigenvector maps (dbMEMs) that significantly explained variation in the allele frequencies.

6.3.6 Larval dispersal simulations

Simulations of larval dispersal with and without the incorporation of biological parameters (i.e temperature-dependency and mortality) were vastly contrasting (Fig. 52). Without any biological parameters considered, larval dispersal across much of the study area was extensive. For example, many particles from Sula Sgeir (northwest Scotland) and Skagerrak were able to travel to the tip of northern Norway in one journey. Moreover, there was considerable mixing in the Celtic and Irish Seas, within the North Sea, and between locations along the western coast of Ireland; however, less mixing was apparent between the western and eastern English Channel. In contrast, when biological parameters were considered, larval dispersal across the study area was notably reduced. In particular, there was very limited dispersal out of northern Scotland and the furthest drifting particles from Skagerrak were only able to travel to southwest Norway. Moreover, dispersal was more limited from all study sites across the seas of Britain and Ireland.

The connectivity matrices generally supported results from the dispersal trajectories, whereby connectivity between sites was much higher with the exclusion of biological (non-bio) parameters compared to the inclusion of biological (bio) parameters (Fig. 53). In the non-bio matrix, asymmetrical connectivity was apparent from western North Sea sites (Brd, Cro, Eye, and Oos) to eastern North Sea (Hel) and Skagerrak sites (Flo, Sin and Kav); this was not the case for the bio matrix in which no asymmetrical connectivity was found across the same spatial scales. Moreover, in both connectivity matrices, Skagerrak sites did not act as a source for other study sites other than for those within Skagerrak, except for a few potential recruits into Sbs and Hel from Flo (Fig. 53). Both matrices indicated that the strongest source-sink relationships were predominantly between the same sites, or between neighbouring sites situated spatially very closer to each other. However, in the bio matrix, there was no intra or inter-site connectivity in three study sites (Ork, She and Sul), all of which are located in colder waters around the north of Scotland.

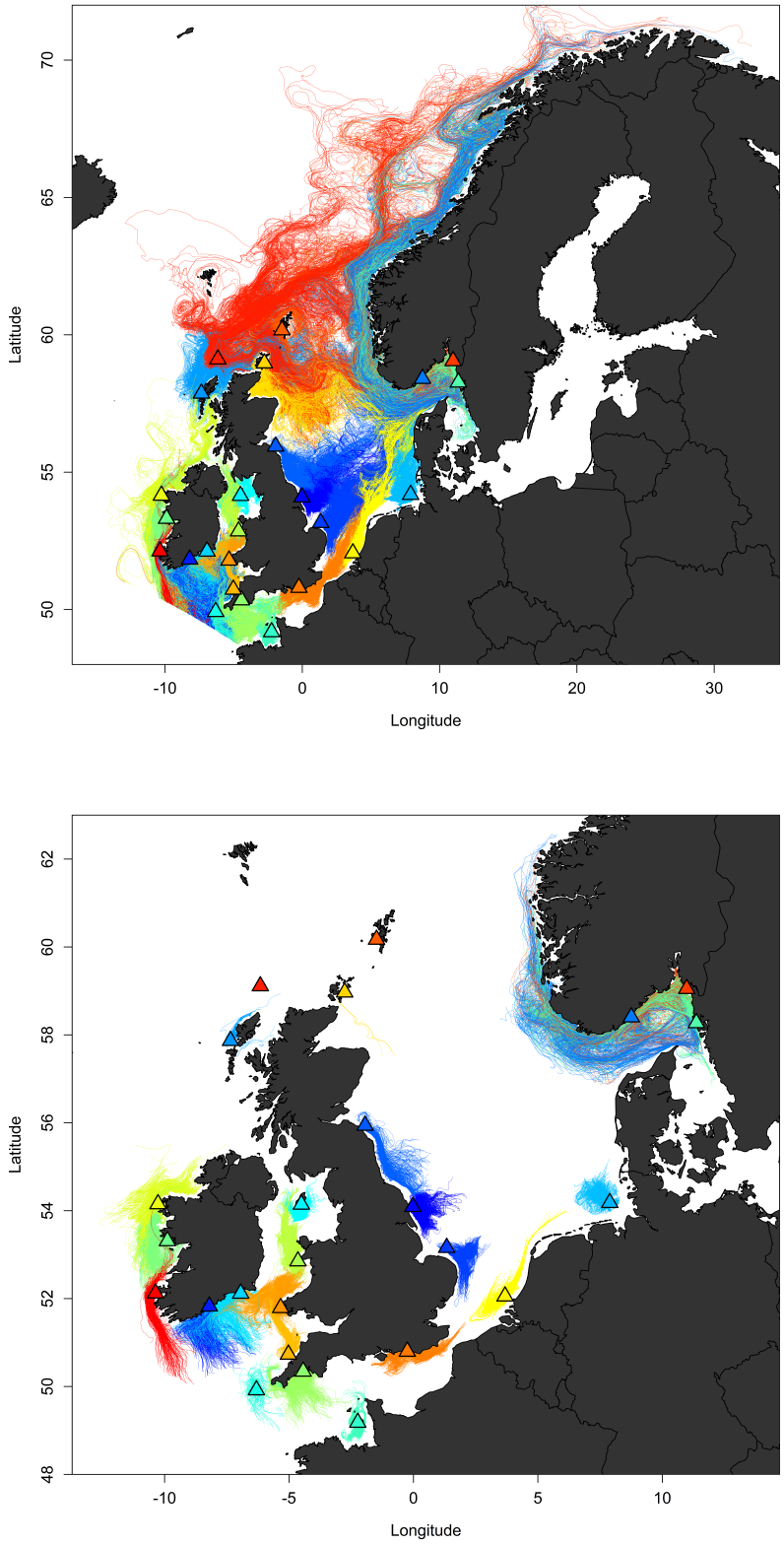


Figure 52: Larval dispersal trajectories with the exclusion of biological parameters (top) and the inclusion of temperature-dependent PLD and mortality (bottom).

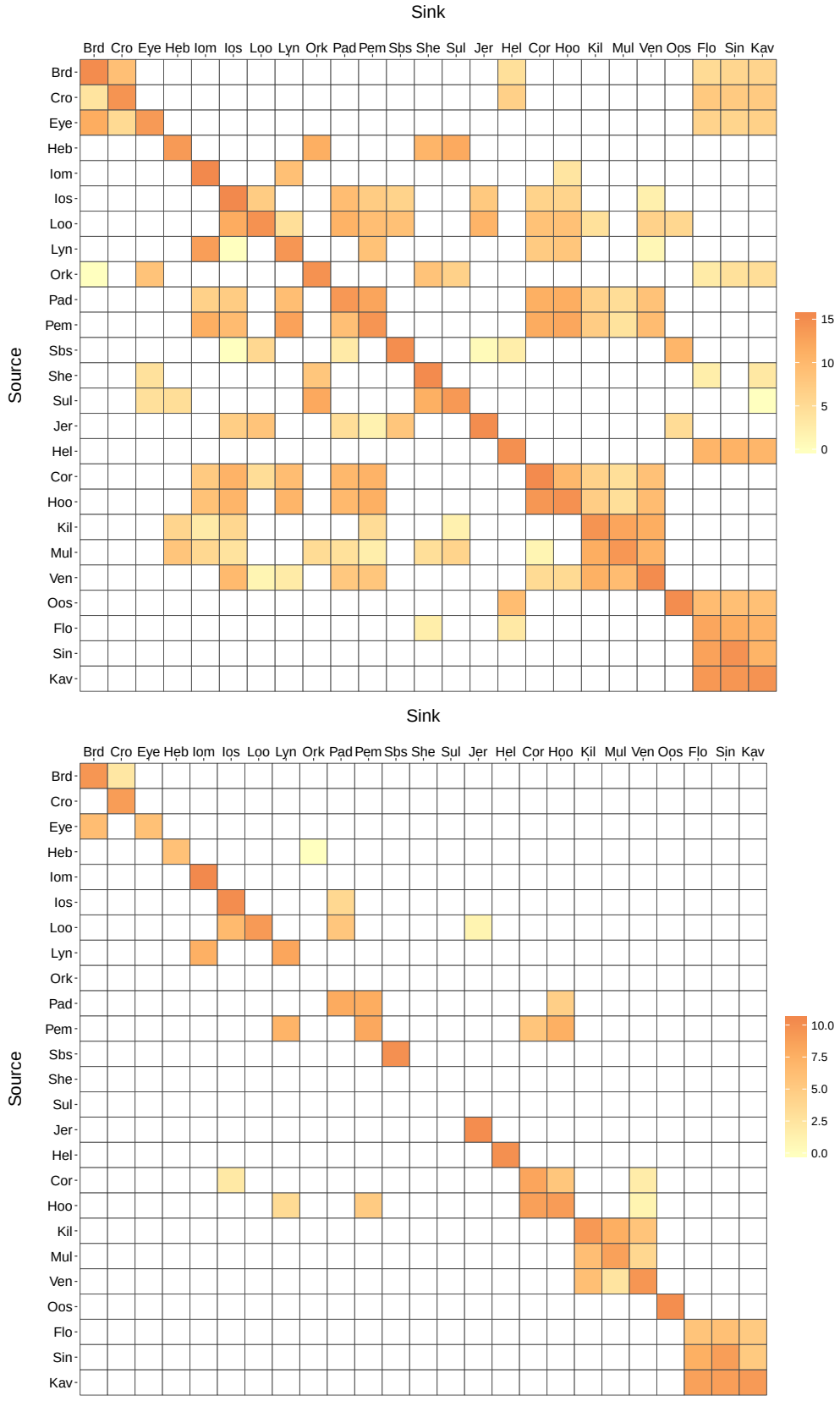


Figure 53: Larval dispersal connectivity matrices with the exclusion of biological parameters (top) and the inclusion of temperature-dependent PLD and mortality (bottom).

6.3.7 Individual assignment

Assigning individuals to their basin of origin (Atlantic or Mediterranean) using the baseline data was extremely accurate, ranging from 98-100 % depending on the proportion of individuals used in the training dataset and the number of loci used for the assignment tests (Fig. 54). When all 86 SNPs were used, the model successfully predicted the basin of origin of the test individuals at 100 % accuracy, regardless of the proportion of individuals used in the training set. This was then tested on the unknown dataset and the SVM and LDA functions both correctly predicted the basin of origin for all 76 individuals. In comparison, assigning individuals back to their location of origin was highly inaccurate, with an overall assignment accuracy of $\sim$1 % (Fig. 54). However, assigning individuals to one of three regional groups (Skagerrak, North Sea, and remaining Atlantic sites) was much more accurate, with an overall assignment accuracy of $\sim$78 % (Fig. 54). The assignment tests suggested that assigning the remaining Atlantic sites back to the Atlantic group was approximately 94 % accurate using all 86 SNPs; assigning Scandinavian (i.e. Skagerrak) sites back to the Skagerrak group was approximately 68 % accurate using all 86 SNPs. In contrast, the assignment tests showed that North Sea sites could not be assigned back to a North Sea group - they were either assigned to the Atlantic group (78 %) or the Skagerrak group (22 %).

6.4 Discussion

This study used a small panel of informative SNPs to investigate population genetic patterns of *H. gammarus* at sites sampled across most of its contemporary geographical range. Although this SNP panel composes loci that were chosen to capture genetic differentiation at different spatial scales, the overall panel includes loci across the spectrum of the differentiation statistics (Fig. 37), thereby permitting the analysis of both loci potentially under high selection or drift and loci which are generally homogenous across certain spatial scales. Sampling density reflected that of lobster abundance, being most concentrated around Britain and Ireland, where approximately 75 % of the species total catch resides (FAO 2018). The patterns of population genetic structure detected in this study supports the general assertion of previous studies that regional lobster stocks are relatively well-connected. However, this study provided much higher resolution of population structure than has been previously achieved, which has enabled the detection of fine-scale differences between sampling sites, particularly in the northeast Atlantic. Crucially, the higher power offered by SNP markers, combined with biophysical modelling, allowed the potential drivers of these patterns to be explored with much

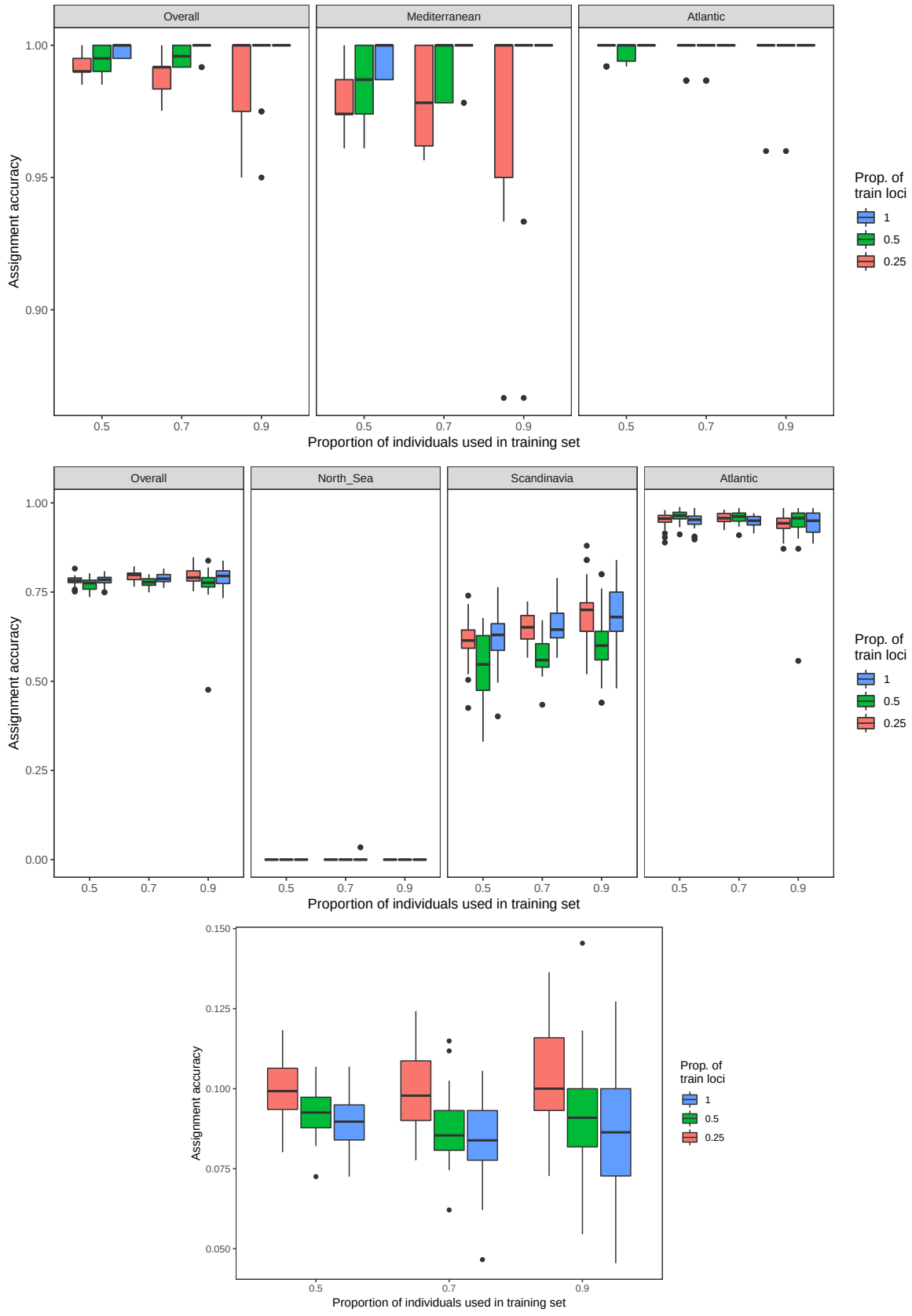


Figure 54: European lobster individual assignment: basin-wide analysis (top), regional analysis (middle), and sampling location analysis (bottom).

greater precision. Genetic and modelling analysis suggested that patterns of connectivity are driven, at least in part, by spatial distances between sites (IBD) which implies that *H. gammarus* larval dispersal likely accords with a stepping-stone model, a finding consistent with a recent study that used 14 microsatellite markers to explore population structure (Ellis et al. 2017). Individual assignment was shown to be highly accurate at a basin-wide scale and reasonably accurate at a regional scale, although marker power and/or underlying differentiation was not sufficient to accurately assign individuals to their sampling location of origin. The findings from this study should be valuable to fisheries management, particularly for providing estimates of stock connectivity and delineating management units (Reiss et al. 2009), and for defining the spatial bounds of ecologically responsible hatchery stock enhancement programmes.

6.4.1 Atlantic-Mediterranean transition

Population structure analyses in this study found three distinct genetic groups, organised into sites from the northeast Atlantic, the middle Mediterranean, and the Aegean Sea (the eastern Mediterranean). Divergence was particularly strong between sites originating from the northeast Atlantic and the Mediterranean, a pattern that was detected in two previous studies that used mtDNA RFLPs (Triantafyllidis et al. 2005) and allozymes (Jorstad et al. 2005). A deep partition between the Atlantic and the Mediterranean basins has been found in previous studies for a diverse array of marine organisms including the European spiny lobster (mtDNA and microsatellites, Palero et al. 2008; 2011), other decapods such as crabs (mtDNA, Roman & Palumbi 2004; García-Merchán et al. 2012) and shrimp (mtDNA, Reuschel et al. 2010), molluscs (microsatellites, Pérez-Losada et al. 2002; mtDNA and allozymes, Sá-Pinto et al. 2012), arrow worms (mtDNA and microsatellites, Peijnenburg et al. 2006), seahorses (SNPs, Riquet et al. 2017), demersal fish (mtDNA and allozymes, Bargelloni et al. 2003, 2005; mtDNA and microsatellites, Pita et al. 2010), and deep-sea fish (mtDNA, Charrier et al. 2006).

The majority of these studies ascribe this partition to restricted gene flow between the Atlantic and Mediterranean basins, possibly due to IBD and/or an oceanographic barrier to connectivity. For example, Castilho et al. (2017) found that genetic patterns in the intertidal peacock blenny (*Salaria pavo*) were likely explained by a combination of IBD and asymmetrical gene flow from the Mediterranean to the Atlantic. Conversely, a break in connectivity between the Atlantic and Mediterranean basins has been suggested to occur at the Almeria-Oran front (e.g. Patarnello et al. 2007) and/or the Strait of Gibraltar (e.g.

García-Merchán et al. 2012). For instance, Reuschel et al. (2010) reported a distinct phylogeographic break in the pan-European littoral prawn (*Palaemon elegans*) across the Atlantic-Mediterranean boundary, a finding the authors linked to reduced larval dispersal across the Almeria-Oran front, which is a front formed by the convergence of two distinct water masses and moderated by the currents of the Eastern Alboran Gyre (Tintore et al. 1988). Aside from this oceanographic barrier, historical vicariant events potentially associated with the Strait of Gibraltar have also been suggested to shape genetic patterns of marine taxa across this boundary. For example, during the last glacial maxima, fluctuations in sea levels (up to 120 m below present-day levels) periodically reduced the width and depth of the Gibraltar Strait (Rohling et al. 1998), potentially resulting in recurrent impediments to dispersal across this boundary (Charrier et al. 2006). In addition, some studies have proposed a scenario of secondary contact across this transition zone, possibly due to the re-opening of the Gibraltar Strait some 5.33 Ma following the theorised Zanclean flood (Pérez-Losada et al. 2002) or due to vicariance during Pleistocene glaciations (Taboada & Pérez-Portela 2016), in which secondary introgression occurs between previously isolated and divergent allopatric populations (Bierne et al. 2013). This has been supported by studies that have reported: (i) clinal changes in allele frequencies across the transition zone (located generally west or east of Gibraltar) in several species (e.g. cuttlefish, Pérez-Losada et al. 2002; seagrass, Alberto et al. 2008; seahorses, Riquot et al. 2017); and (ii) the discovery of two distinct Atlantic-Mediterranean clades, of which one clade co-occurs in both basins (e.g. swordfish, Alvarado-Bremer et al. 2005; brittle stars, Taboada et al. 2016).

The results of this study suggests that a neutral pattern of IBD likely explains some of the structure in *H. gammarus* between the Atlantic and Mediterranean basins, supported by DAPC, STRUCTURE and IBD analyses with neutral SNPs. It is also possible that the Almeria-Oran front or the Strait of Gibraltar have contributed to this pattern, but a lack of sampling in the western Mediterranean means it is difficult to make inferences about the potential role of these putative barriers. Moreover, although the genetic cline (Fig. 46) from the eastern Mediterranean to northwest Spain potentially supports a scenario of secondary contact, sites from the southern Atlantic and western Mediterranean are needed to fully explore this hypothesis. Alternatively, analyses with outlier SNPs, which are putatively non-neutral, revealed that some of the basin-wide differentiation detected is explained by divergence at SNPs potentially under the influence of divergent selection. This may suggest that lobsters from these two basins are potentially locally adapted to specific environmental conditions, of which sea temperature and

salinity are possible selective factors as they have been shown to drive adaptive divergence among populations of numerous marine invertebrate species (Sanford & Kelly 2011; Dalongeville et al. 2018). For example, a temperature gradient (-1°C to 26°C) exists across the northwest Atlantic distribution of *H. americanus*, and a recent study found evidence for thermal adaptation in this species (Benestan et al. 2016b). A similarly extensive thermal gradient exists across the range of *H. gammarus* populations sampled in this study –from the Aegean Sea (23.5-26.4°C in August) to Skagerrak (1.1-6.3°C in March)– so future research may investigate temperature as a potential driver of adaptive variation. A recent study on seahorses suggested that a gradient of introgression between Atlantic and Mediterranean lagoons is likely driven by parallel outlier loci, possibly from adaptive introgression or because of a shared history of divergence retained at outlier loci against secondary gene flow (Riquet et al. 2017); based on the patterns at outlier SNPs in this study, this could also be a potential explanation for the Atlantic-Mediterranean partition in *H. gammarus*.

6.4.2 Differentiation within the Mediterranean

In addition to the strong differentiation observed between the Atlantic and Mediterranean basins, this study detected differentiation (albeit slightly weaker) in *H. gammarus* within the Mediterranean Sea, separated into the middle (Sardinia and Lazio) and eastern Mediterranean (Aegean Sea). A similar pattern was found in a previous study of *H. gammarus*, whereby samples from the Aegean Sea were differentiated from one sample in the Adriatic Sea (mid-Mediterranean) and one sample from the Columbretes Islands (western Mediterranean) (Triantafyllidis et al. 2005). In the present study, the stronger differentiation observed at neutral SNPs infers restrictions to gene flow; the Mantel test using only Mediterranean sites (although non-significant) (Fig. 50d) indicated that IBD may be a key driver of this pattern, which suggests that *H. gammarus* larval dispersal between the middle and eastern Mediterranean follows a stepping-stone model of connectivity. In addition, as clinal patterns were found in the STRUCTURE analysis using outlier SNPs, this could also suggest that secondary contact may explain some of the differentiation observed, potentially from adaptive introgression (Riquet et al. 2017). In any case, the evidence for divergence via drift implies that, despite the generally high N_e expected in marine invertebrates, N_e across these spatial scales is not sufficiently large to mitigate drift. This may be due to the present-day lower abundance and patchy distribution of *H. gammarus* in the Mediterranean (resulting from past over-exploitation), which is also supported by the overall lower genetic diversity

(possibly due to bottlenecks) found in the Mediterranean sites in this study.

6.4.3 Northeast Atlantic connectivity

In the northeast Atlantic, there appeared to be a genetic cline in the datasets that used all SNPs and outlier SNPs (and in the neutral SNPs using STRUCTURE K=3), starting from the most southerly sampling site (Vigo, northwest Spain) to the most northerly sampling sites (Flo-Gul-Kav-Lys, Skagerrak). Interestingly, however, this pattern was considerably weaker in the DAPC using neutral SNPs; instead, most northeast Atlantic sites generally clustered together, with the exception of Oosterschelde which was partially differentiated from the main cluster of Atlantic samples.

Mantel tests conducted with only Atlantic sites (Fig. 50b) and neutral SNPs found that geographical distances explain 17 % of the variation in the F_{st} dissimilarity matrix, although this increased to 49 % when Oosterschelde was removed (Fig. 50c). This suggests that IBD likely explains a considerable component of the genetic cline observed, which is also supported by the biophysical modelling as the dispersal trajectories and connectivity matrices generally imply that seascape hydrology (i.e. ocean currents) facilitates a stepping-stone model of connectivity between the study sites (Fig. 52, 53). Nevertheless, assuming that IBD is indeed involved in driving some of this genetic cline, over 50 % of the variation in the F_{st} dissimilarity matrix (and 72 % in the RDA using all sites) remains unexplained, which suggests other processes are also responsible for shaping this genetic cline.

As alluded to previously, the most commonly proposed causes of clinal patterns in allele frequencies are (i) IBD caused by neutral drift, (ii) selection across an environment gradient, and (iii) secondary contact and introgression between previously isolated and genetically divergent populations (Pérez-Losada et al. 2002). In this study, Mantel tests with neutral SNPs provided support for IBD. However, fine-scale local adaptation across an environmental gradient (e.g. temperature), similar to the explanation proposed for the differentiation between the Atlantic and Mediterranean basins, cannot be ruled out as a driver of this pattern. This is supported by analysis of the 15 outlier loci, which showed a very clear genetic cline in both the DAPC and STRUCTURE analyses. Yet, without the incorporation of sea temperature data, and further genomic resources that have reliable gene annotations (e.g. whole genome or transcriptome), local adaptation to sea temperatures and the mechanisms behind this process are speculative. Alternatively, this clinal pattern could be explained by secondary contact following range expansions from refugia after the Last Glacial Maximum (or more ancient

glaciations in the Pleistocene). Putative refugia in the northeast Atlantic have been proposed in western France, the Iberian Peninsula, and west-southwest Britain and Ireland (Maggs et al. 2008; Finnegan et al. 2013; Jenkins et al. 2018a), evidenced by the high levels of genetic diversity found in populations inhabiting these areas (Provan & Bennett 2008). Given that the highest genetic diversity in this study was found in the Bay of Biscay (Île de Ré) and northwest Spain (Vigo), it is possible these two sites served as glacial refugia, which preceded secondary introgression of northward dispersers after the ice retreated.

Mantel tests in this study also revealed a reduced correlation (and a substantial decrease in significance) when lobsters from Oosterschelde were excluded from the analysis. Moreover, using neutral SNPs, there was evidence that the Oosterschelde sample was genetically differentiated from lobsters at all other Atlantic sites. This pattern has been reported in a previous study on *H. gammarus* using mtDNA RFLPs (Triantafyllidis et al. 2005); the authors attributed this differentiation in Oosterschelde to a combination of past bottlenecks and a lack of immigration. In 1962-1963, harsh winters brought extremely low water temperatures and salinities (from high localised river discharges) to Oosterschelde, causing mass mortality of lobsters and many other marine organisms (Triantafyllidis et al. 2005). Moreover, construction of dams over the last century has created a semi-enclosed area virtually isolated from the North Sea (Triantafyllidis et al. 2005). The differentiation and lower diversity of lobsters from Oosterschelde found in this SNP study are in line with the conclusions by Triantafyllidis et al. (2005), that is, this pattern is likely a result of drift caused by past bottlenecks and a barrier to gene flow over the last century.

At much finer spatial scales (i.e. within and between neighbouring seas), and taking the genetic cline into account, the results from this study suggest high genetic connectivity between certain sites sampled in the northeast Atlantic. For example, high gene flow was apparent within and between sites situated in the English Channel, the Celtic and Irish seas, the coast of western Ireland and up to northern Scotland (up to distances of 1,400 km); this was supported by the relatively low pairwise differentiation indices (F_{st} and D) and the population structure analyses. This was also supported, but to a lesser extent, by the biological modelling simulations. Although the non-bio dispersal trajectories and connectivity matrices indicated extensive dispersal potential across these scales, this is not likely to be realistic, or as representative of real-world transience as the bio projections of dispersal and connectivity. When biologically realistic dynamic parameters were factored into the model, the results showed significantly less dispersal potential, with practically no dispersal from sites north of Scotland (Ork,

Sul and She), where larvae encountered a median temperature of $<14^{o}$C for more than half of their drift time which meant they were considered dead. However, these areas host lobster stocks which have long supported intensive commercial fisheries, into which recruitment must occur somehow. Quinn et al. (2013) showed that, when reared in cold water (10^{o}C), larval development times among *H. americanus* clutches sourced from females at the northern extent of the range were significantly reduced compared to those of clutches sourced from more southerly latitudes. The authors interpreted this, and results showing the reverse trend when cold-water larvae were reared at higher temperatures, as a signal of local adaptation to their environment, and it seems plausible that *H. gammarus* larvae originating from the northerly areas sampled in this study have undergone similar adaptations to their thermal niche. If this were the case, then the dispersal potential of larvae from northerly areas may well have been underestimated by our biological parameters, which were all based on the development of larvae sourced from females inhabiting warmer waters to the south (i.e. Helgoland and southwest England) (Schmalenbach & Franke 2010).

High genetic connectivity at similar fine-scales were also found in other regions; for example, between sites within the Skagerrak region, between sites in the middle Mediterranean, and between sites within the Aegean Sea. For Skagerrak, this was supported by both (bio and non-bio) simulations of larval dispersal, which implies that connectivity between sites in Skagerrak can occur in a single dispersive event at distances of up to 168 km (Flo-Sin). Moreover, these simulations suggest that larvae can disperse to sites in western Norway; though, previous studies have found northern Norway populations to be genetically distinct (Triantafyllidis et al. 2005; Jørstad et al. 2005), which indicates that sites in northern Norway may be more isolated than the simulations in this study suggest. The genetic isolation of northern Norway could not be tested in this SNP study; nevertheless, based on these previous studies, it appears that gene flow between northern Norway and other sites in the northeast Atlantic may be restricted or that selective forces occurring at the northern range edge may be mitigating the homogenising effect of gene flow. In the Mediterranean, the high genetic similarity of sites (i) within the middle Mediterranean and (ii) within the Aegean Sea, suggests that genetic connectivity can potentially occur up to distances of 294 km (between Sardinia and Lazio) and 280 km (between Alexandroupoli and Thermaikos Bay), respectively. However, because the spatial extent of the ocean model did not facilitate simulations of larval dispersal across the Mediterranean, it is difficult to ascertain whether these patterns are due to high gene flow or due to high N_e (or due to both). Alternatively, the high genetic similarity between sampling sites that are

situated closely together (e.g. sites within Skagerrak) may suggest that these sites are a single panmictic population.

6.4.4 Implications for fisheries management and stock enhancement

For managing fisheries, it is important to identify stock structure and connectivity to ensure that the spatial implementation of management is commensurate with that of biological population units (Reiss et al. 2009), and to pinpoint populations that may contribute colonisers to overfished or depleted stocks (Da Silva et al. 2015). The latter typically relies on estimating demographic connectivity, which is difficult to measure, particularly with genetic data alone (Lowe & Allendorf 2010). In the northeast Atlantic, analysis of neutral SNPs and biophysical modelling from this study provided evidence for a stepping-stone model of connectivity, in which larvae have the potential to disperse up to 300 km from some coastal sites of the British Isles (Fig. 52). This implies that site-specific recruitment may not always come from local sources, but from adjacent local or regional sources. Thus, if multiple adjacent (LFU) stocks across Britain and Ireland collapsed simultaneously, this could have profound ramifications for stock recovery and local fisheries. Recent research has found that temporary closures or prohibiting fishing in MPAs offers some respite to lobster populations (Moland et al. 2013; Roach et al. 2018), which may be a viable management option for lobster fisheries going forward to prevent over-exploitation. However, safeguarding lobster stocks in one area via temporary closures or MPAs (e.g. Isles of Scilly) may inadvertently benefit local and regional areas adjacent to the protected area, due to an increase in larval subsidy, rather than the area being protected. Therefore, fisheries management should acknowledge these patterns of dispersal and recruitment when considering which source populations to protect during temporary closures and for designing holistic management programmes.

For lobster hatcheries, knowledge of stock structure is crucial to ensure that reared juveniles, which are usually reared from the egg clutches of wild females originating from the local area (Ellis et al. 2015c), demonstrate evolutionary compatibility with the targeted population being stocked (Ward 2006). Overall, the genetic profiles observed in this study using neutral and outlier SNPs suggests that stocking of a target population should ideally be implemented with juveniles whose parents originate from the same geographical area. Furthermore, the use of broodstock originating from the northeast Atlantic to stock target populations in the Mediterranean, or vice versa, is to be highly discouraged; this also applies to broodstock originating from the middle Mediterranean with the aim to stock target populations in the Aegean Sea (and vice versa). This is because of the potential to

introduce maladapted traits into the target population that could also proliferate to neighbouring populations (Araki et al. 2007); this has the most deleterious consequences in stocks that are depleted or are highly adapted to local conditions (Lorenzen et al. 2012). In this study, a significant excess of heterozygotes were found at several sites, which may indicate the presence of heterozygote advantage at some loci (Sellis et al. 2011), or outbreeding depression, which can be caused by stocking a target population with individuals adapted to a completely different environment (Frankham et al. 2011). For example, three of the five Scandinavian sites in Skagerrak showed this pattern, of which Kvitsoy (southwest Norway) $\sim$225 km away from Flodevigen has seen quite extensive stocking during 1990 and 1994 (Agnalt et al. 2004; Ellis et al. 2015c). Although this result is potentially evidence for outbreeding depression at these Scandinavian sites, most of the broodstock used in the stock enhancement programme were reported to originate from Kvitsoy (Agnalt et al. 2004).

Individual assignment using genetic techniques has been shown to be a potentially useful tool for determining the origin of fished individuals and for tackling IUU fishing (Martinsohn & Ogden 2009; Nielsen et al. 2012; Bernatchez et al. 2017). However, the power of the molecular markers employed is highly sensitive to the degree of genetic differentiation between sites (Christie et al. 2017). This study demonstrated that a panel of 86 SNPs has adequate power to assign lobsters to either the Atlantic or Mediterranean basin at 100 % accuracy. This may have useful applications for management authorities, particularly those responsible for Mediterranean coasts, who wish to find out whether an individual lobster has been imported from somewhere in the northeast Atlantic or has been locally caught in the Mediterranean. Moreover, managers could test for the introduction of Atlantic lobsters into the Mediterranean via escapees or incidental larval release (i.e. from storage facilities on the coast), or from inadvisable attempts to boost stocks through the intentional release of larvae, juveniles or adult *H. gammarus* that originate from Atlantic populations. Unfortunately, it is not currently possible to assign lobsters back to their location of origin using this SNP panel; however, it may be possible to assign lobsters with some confidence to a geographical region (e.g. Skagerrak). The power and accuracy of these assignment tests at local and regional scales could be improved by incorporating more SNP markers and by using software that take patterns of IBD into account (e.g. Guillot et al. 2016; Drinan et al. 2018).

6.4.5 Limitations and future research

The drivers of some of the genetic patterns found in this study, such as secondary contact, could be further explored with samples from sites in southern European Atlantic waters (i.e. southern Portugal and Spain) and Atlantic Morocco, and from sites in the western Mediterranean (i.e. Alboran Sea). Moreover, although analysis of outlier SNPs in this study was linked to differentiation at the southern range limit (Aegean Sea), samples from northern Norway would be advantageous to explore the potential for local adaptation to sea temperatures at both the northern and the southern range limit of *H. gammarus*. Future research could also collate sea temperature (and salinity) data to test whether these data explain any of the variation in allele frequencies; such an approach was implemented in the American lobster to explore thermal adaptation (Benestan et al. 2016b).

Analysis to detect outlier SNPs in this study was stringent; yet, as the panel was composed of a subset of informative SNPs derived from RADseq data, there is a possibility some of these outlier SNPs are false-positives. However, 13 out of 15 of the SNPs classified as outliers in this study were also classed as outliers in the raw RADseq dataset (Chapter 5), so it is likely that these 13 SNPs are genuine outlier markers potentially under divergent selection. Moreover, although two outliers were not detected in the original RADseq dataset, both SNPs were still classed as outliers in this study because they could represent adaptive loci associated with sampling sites which were included in this study but not the RADseq study. All 15 SNPs could be further validated as outliers by aligning the raw RAD-tag loci to the American lobster or European lobster transcriptome, although at present only the American lobster transcriptome is publicly available. Additionally, the population allele frequency of one allele for each of the 15 outlier SNPs was visualised (Fig. 49). This showed some interesting differences in the frequency of the allele investigated, with several SNPs completely fixed for one allele in some sampling sites from the Atlantic, the middle Mediterranean or the eastern Mediterranean. These loci could act as diagnostic SNPs that differentiate populations from the Atlantic, the middle Mediterranean and the eastern Mediterranean; moreover, these SNP loci may also be useful for future research to estimate admixture and migration.

Interpreting genetic data using STRUCTURE analysis in the presence of IBD or hierarchical structuring can be challenging (Frantz et al. 2009; Gilbert 2016). The STRUCTURE admixture model assumes that each of the K ancestral populations existed at some point in the past and that modern individuals were produced by recent mixing (or no mixing) of these ancestral populations (Lawson et al. 2018).

Incorrect inferences of *K* may arise when hierarchical levels of population structure exist (subpopulations within populations) and there is uneven sampling across these levels (Puechmaille 2016). However, the sampling regime in this study was robust, with regularly distributed samping locations in the northeast Atlantic and with the majority of sampling sites containing approximately 35-37 individuals; this mitigates the potential caveat of uneven sampling. This study also found strong spatial autocorrelation across the study area, which may limit the ability of STRUCTURE to accurately delimit clusters and assign individuals to ancestral populations. For example, Frantz et al. (2009) showed that Bayesian clustering methods can overestimate genetic structure when analysing genetic data characterised by IBD (or clines of genetic variation across a landscape). This is a potential limitation of the underlying model when applied to these scenarios, in which the inferred value of *K* and the corresponding allele frequencies in each cluster can sometimes be arbitrary, a limitation acknowledged by the authors of STRUCTURE (Perez et al. 2018). Although these caveats should be considered in this SNP study, the STRUCTURE results were consistent with other population structure analyses (e.g. genetic indices and DAPC). However, as IBD appears to be a characteristic of this dataset, future analyses such as PCA and tree-building based on unbiased genetic distances may be more appropriate to avoid such caveats.

6.4.6 Conclusion

In conclusion, the results of this SNP study indicate that several factors have potentially contributed to shaping the genetic structure of *H. gammarus* across its northeast Atlantic and Mediterranean distribution, including IBD, oceanographic barriers and secondary contact after a period of allopatric isolation. How these patterns, and lobsters in general, will respond to projected increases in sea temperature with global climate change is uncertain. Indeed, as the PLD of *H. gammarus* is temperature-dependent (Schmalenbach & Franke 2010), elevated sea temperatures could see future range expansions northward, perhaps into the Faroe Islands and Iceland; such a pattern of expansion has been suggested in the green crab (*Carcinus maenas*) (Roman & Palumbi 2004), whereby Shetland has potentially acted as a stepping-stone for range expansion. In contrast, warming temperatures may act to decrease the PLD (possibly reducing dispersal distance and connectivity) and potentially increase larval mortality in the eastern Mediterranean where sea temperatures across the range of *H. gammarus* are highest (Schmalenbach & Franke 2010). Consequently, managing and monitoring

lobster stocks effectively, and preserving the distribution of genetic diversity across the range of *H. gammarus*, will likely be critical in ensuring the future persistence and sustainability of lobster populations and the fisheries they support.

Chapter 7: General discussion

This thesis has addressed three research components: (i) exploration of comparative phylogeography in a number of diverse taxa across the northeast Atlantic using meta-analysis; (ii) proposed criteria for selecting candidate species to assess connectivity between MPAs; and (iii) generated novel population genetic data for two benthic marine organisms, the pink sea fan and the European lobster, which has provided novel evidence to address connectivity between MPAs. In this final chapter, the results of this thesis are summarised, followed by a discussion on the translation of genetic data into policy and speculation on how future developments may enhance studies of marine connectivity.

7.1 Summary and addressing research hypotheses

The meta-analysis of coastal marine taxa (Chapter 2) represents one of the first meta-data studies to investigate comparative phylogeography in the northeast Atlantic; the other meta-analysis study in the north Atlantic explored the possibility of distinguishing signatures of periglacial refugia from southern refugia (Maggs et al. 2008). The inclusion of many diverse taxa in the present meta-analysis allowed two main hypotheses to be tested: (i) whether there are common phylogeographic breaks among taxa; and (ii) whether demography was constant or variable among coastal species during the Pleistocene glaciations, and in particular during and after the LGM. The results showed that marine taxa in the northeast Atlantic show a mixture of contemporary genealogical structure and that patterns of phylogeography are discordant, which supports the null hypothesis of contrasting patterns of phylogeography. This finding generally accords with similar meta-analyses across other seas/oceans, which also report contrasting patterns of phylogeography across the southeast coast of the United States of America (Pelc et al. 2009), the northeast Pacific (Marko et al. 2010), and the northwest Pacific (Ni et al. 2014). The results in the meta-analysis from this thesis also indicated that population expansions were common in northeast Atlantic coastal taxa and were mostly linked to the LGM, providing evidence to reject the null hypothesis and accept the alternative hypothesis; this agrees with the findings for five species studied by Marko et al. (2010) but is contrasting to the findings of Ni et al. (2014) in which the authors reported that most population expansions pre-dated the LGM.

Selection of candidate species for assessing genetic connectivity between

MPAs (Chapter 3) was tailored towards species whose distribution spanned MPA boundaries designated in UK waters. However, the criteria proposed to select taxa in this thesis could easily be applied to assessments of MPA network connectivity in other seas and oceans around the world. Although no hypotheses were tested in this chapter, this research component was essential to this thesis because it facilitated the selection of appropriate taxa for addressing MPA connectivity across the British network, particularly across the English and Welsh network.

To explore the population genetics and connectivity of the pink sea fan and the European lobster, genomic techniques that utilise NGS technology were employed; this is the first time RADseq or a similar approach has been used to isolate genome-wide SNPs in either of these two non-model organisms. For pink sea fan, there was distinct population structure using both microsatellite and SNP markers, organised into sites from Britain-France, southern Portugal and northwest Ireland, with some evidence for weak differentiation between samples from Britain and France. Therefore, the null hypothesis can be rejected and the second alternative hypothesis (H_{1B}) can be accepted. In addition, the results suggested that the MPA network in southwest Britain is likely sufficient for maintaining genetic connectivity in pink sea fans; thus, the null hypothesis can also be rejected and the first alternative hypothesis (H_{1A}) can be accepted. For European lobster, there was strong differentiation between lobsters from the northeast Atlantic, the middle Mediterranean and the eastern Mediterranean; therefore, the null hypothesis of no population structure can be rejected. In the northeast Atlantic, a genetic cline from northwest Spain to Skagerrak was evident, which could be explained by one or more processes such as IBD, fine-scale local adaptation and secondary contact. Consequently, depending on which sites are being compared, this provides potential support for the first alternative hypothesis (i.e. between sites spatially close together) and for the second alternative hypothesis (i.e. between sites spatially further away). Moreover, further analysis of neutral SNPs suggests that Oosterschelde is genetically isolated from other sites in the northeast Atlantic, which supports a hypothesis of population structure driven by reduced gene flow and subsequent drift. The connectivity matrix based on the simulation model that incorporated biological parameters suggested that larval dispersal can potentially occur at spatial scales up to hundreds of kilometres, which was supported by overall genetic homogeneity at these scales. Therefore, based on the recommended placement of a MPA every 80 km or less by Roberts et al. (2010), the results for European lobster from this study suggest that the MPA network across England and Wales is likely sufficient to maintain genetic connectivity, but potentially also demographic connectivity.

7.2 Assessing connectivity between MPAs: which taxa?

Connectivity is identified as one of five key principles for designing an ecologically coherent network of MPAs in European waters (OSPAR Commission 2017). Yet, without the ability to demonstrate connectivity, it is impossible to be certain that sites designated within a MPA network do in fact constitute a network, when they may be in reality a set of unlinked habitats and associated species assemblages. In this thesis, a set of criteria were developed to pinpoint certain taxa that may represent ideal surrogates for empirically assessing connectivity between and outside of MPA boundaries. Particular attention was afforded to taxa which may fulfil the criteria of an umbrella, keystone or flagship species (Simberloff 1998; Kalinkat et al. 2017). For example, connectivity patterns observed in one species may be representative of other species with similar biology and dispersal capacity (i.e. similar PLDs); consequently, when species assemblages across an area of interest are well documented, designing a network based on known connectivity patterns from one species may extrapolate benefits to other organisms in the community (Marti-Puig et al. 2013). As the biology and dispersal of European lobster is well understood, such an approach could be used to fill in gaps in the current MPA network around Britain. However, species whose dispersal traits are poorly understood should not necessarily be dismissed, particularly if the species has high local or national conservation priority. These species may be advantageous for driving MPA proposals and designations that encompass populations of these species; pink sea fans in southwest Britain is such an example, whereby several MCZs have been specially designated to protect populations of pink sea fan, some of which may have added benefits to other (non-protected) species in the MCZs with similar life histories and dispersal traits (e.g. dead man's fingers, Holland et al. 2017).

7.3 Translating genetic data into policy

Content discussed in this section is based on a paper published in the journal *Marine Policy* (section 3, Jenkins & Stevens 2018). A central premise of this thesis is to use the novel genetic data generated to inform practitioners in the fields of marine protected area design and fisheries management. However, translating primary research into the language and terminology required by policymakers and conservation managers is not a trivial task. Often, it may be more beneficial to present a few points that represent the key findings of a study, while trying to avoid unnecessary technical jargon, which could lead to misinterpretation or confusion. Several papers have discussed the challenges of translating genetic data to inform

management and have asserted the importance of strong collaboration and communication between scientists and practitioners (Laikre et al. 2010; Gordon et al. 2014; Shafer et al. 2015; Galla et al. 2016; Garner et al. 2016; Hogg et al. 2017; Taylor et al. 2017). Some of the reasons put forward for the avoidance of genetic data in fisheries management include a lack of understanding of the potential value of genetic data, the assumption that genetic studies are expensive, and the suggestions that other data types are significantly more important than genetic information in management decisions (Bernatchez et al. 2017). One feature of genetic data is that they cannot be seen or measured without the use of specialist molecular techniques, meaning it can sometimes be difficult to articulate the level of variation and the importance of genetic diversity to non-scientists (Laikre et al. 2010). Moreover, in cases where research is carried out by non-academic bodies, these institutions often have little incentive to publish, or have internal deadlines or political/legal constraints that may delay scientific publication, so the findings may not be widely disseminated (Garner et al. 2016).

However, while some barriers to the dissemination of genetic research exist, there are examples across various taxa and systems where genetic data have successfully informed policy and conservation, and have led to improved management decisions. This suggests that some barriers to the application of genetic data are starting to be overcome. Some examples include the genetic restoration of Florida panthers (Johnson et al. 2010), the genetic management of salmonids (Baerwald et al. 2011; Habicht et al. 2012), the authenticity and monitoring of seafood in sushi bars (Vandamme et al. 2016), and the traceability of fisheries resources (e.g. FishPopTrace, Martinsohn & Ogden 2009).

Yet, while there are a myriad of studies documenting the spatial genetic structure and genetic connectivity of benthic marine species, very few of these studies have been directly used as evidence to inform or support MPA designations and/or network connectivity. This may be a consequence of ineffective dissemination of the key findings of research projects, but also likely relates to the availability of data at the time when large-scale MPA projects were commissioned and candidate lists were first drawn-up. Nevertheless, as these data are becoming more available to practitioners, it is crucial that gaps between primary research (i.e. academic researchers) and applied science (i.e. policymakers) are overcome in order to realise the potential of genetic data to inform MPA design and conservation planning (Shafer et al. 2015; Garner et al. 2016).

Genetic data are currently not (to the authors' knowledge) used by managers as evidence to inform MPA designation or network connectivity in England and Wales. Discussions with national agencies suggest that the personnel and infrastructure

are not in place to process, grade and assess the usefulness of spatially relevant genetic data (Jenkins & Stevens 2018). This may explain the lack of genetic data currently used as evidence to support existing MPA designation or to inform new designations around southwest Britain (Jenkins & Stevens 2018). However, genetic data from single-species can provide an estimate of realised connectivity within evolutionary timescales and, combined with biophysical modelling, these data would likely supplement the present methods used to assess network connectivity in Britain (section 1.3.4). Moreover, genetic data can reveal distinct localised genetic diversity, otherwise undetectable using only presence/absence data or modelling, which can be of major importance for identifying populations or areas that should be prioritised for protection (Funk et al. 2012).

The pink sea fan microsatellite study (Holland et al. 2017) and the SNP study in this thesis has the potential to inform and support the designations of MPAs that include *E. verrucosa* as a protected feature. The key finding from these studies which might constitute evidence for MPA manager is that, as it stands, the MPA network in southwest Britain appears to be adequate to maintain genetic connectivity in this protected species. This also appears to be the case for European lobster, for which the MPA network may be able to maintain both genetic and demographic connectivity. The integration of these data in future reviews or monitoring reports would likely serve as another piece of evidence to support the designation of these MPAs and to help demonstrate the ecological coherency of the network in southwest Britain.

Although genetic data have much promise for informing marine conservation and fisheries management, there are some general limitations. Firstly, managers are typically interested in demographic connectivity, which is difficult to quantify using genetic data, unless combined with other data such as population growth rates, reproductive success or biophysical modelling (Lowe & Allendorf 2010; Breusing et al. 2016). Moreover, as few as ten effective immigrants per generation may be sufficient to maintain genetic homogeneity (drift connectivity) (Lowe & Allendorf 2010), meaning that, despite being genetically similar, some populations may have minimal larval exchange (Botsford et al. 2009). However, genetic markers provide insights into realised connectivity, which may be useful for managers who are interested in the contribution of immigrants that survive, reproduce and add to the local gene pool. Secondly, these approaches usually assume that populations are at gene flow-drift equilibrium; deviation from this assumption can lead to over-estimating the amount of contemporary gene flow (Lowe & Allendorf 2010). For example, inferring patterns of connectivity from marine species with overlapping generations or long-life spans (e.g. corals) can be

challenging because genetic profiles can remain essentially unchanged for many decades, even after barriers to gene flow are introduced. Therefore, in some cases, interpretations of genetic homogeneity may represent historical and not contemporary gene flow and, vice versa, distinct population structure may represent historical and not present-day isolation (Hedgecock et al. 2007). This difference in timescales is critical to consider in assessments of connectivity because MPA networks are generally established to protect and maintain present-day and future patterns of biodiversity and connectivity, or to facilitate recovery/restoration to a previous level of abundance and diversity (Jenkins & Stevens 2018).

7.4 Future developments in marine connectivity

For exploring marine connectivity in benthic organisms with a pelagic larval phase there are three areas that may see pronounced developments in the coming decades: population genetics, biophysical modelling, and real-time tracking. Genomics has begun to revolutionise conservation genetics (Allendorf et al. 2010; Funk et al. 2018) and is likely to continue over the next decade; this will likely have benefits for marine connectivity by providing more accurate measures of population genetic parameters (e.g. gene flow) and by providing higher power for assignment approaches. Of course, in the near future we may have complete genomes sequenced for thousands of species (Allendorf et al. 2010; Fuentes-Pardo & Ruzzante 2017), which could allow researchers to directly compare whole genomes of their study species. In addition, future advances in computer power and memory may generate very high-resolution (i.e. $<$50 m) ocean circulation models; this may enable the development of IBMs capable of tracking of larvae at extremely fine-scales. Lastly, although the technology does not yet exist, methods may emerge that allow researchers to track pelagic larvae in real-time from eggs to settlement; however, there is currently no evidence to suggest this is or will be feasible. In any case, the development of molecular techniques and more powerful ocean models is almost certain, and will likely give unprecedented resolution into the marine connectivity of benthic marine organisms.

Appendix

Chapter two

A1: Mutation rates

Mutation rates used for each species are freely available from the supplementary material hosted online by *PeerJ* (https://doi.org/10.7717/peerj.5684).

A2: Haplotype networks for all species

Haplotype networks for all species are freely available from the supplementary material hosted online by *PeerJ* (https://doi.org/10.7717/peerj.5684).

A3: Mismatch graphs for all species / lineages

Mismatch graphs for all species are freely available from the supplementary material hosted online by *PeerJ* (https://doi.org/10.7717/peerj.5684).

Chapter four

A4: Detailed salting-out DNA extraction protocol

Materials

Proteinase K (20 mg/ml)
RNase A (100 mg/ml)
1 % SDS cell lysis buffer (100mM Tris-Cl; 50 mM EDTA; 1 % SDS)
7.5 M ammonium acetate
0.5 M EDTA
Nuclease-free water
100 % isopropanol
70 % ethanol

Equipment

TissueLyser & microbeads
1.5 ml microcentrifuge tubes
Sterile blue roll
Microcentrifuge & vortexer

Protocol

1. Remove sample from preservative and dap on sterile blue roll to remove excess ethanol.
2. Add samples (up to 30 mg) to 1.5 ml microcentrifuge tubes containing a microbead.
3. Homogenise samples by placing in a TissueLyser for 30 seconds at 30 Hz/s (repeat if necessary).

Digestion

4. Add the following to each tube:
    a. 350 µl 1 % SDS cell lysis buffer
    b. 42 µl 0.5 M EDTA
    c. 10 µl proteinase K.

5. Mix by vortexing and incubate at 65^{o}C for 4 hours.
6. Add 2 µl RNase A and incubate on a thermomixer at 37^{o}C for 30 minutes.

Remove proteins and cellular debris

7. Add 140 µl 7.5 M ammonium acetate to each tube. Mix by vortexing. Incubate at 4^{o}C for 10 minutes.
8. Centrifuge at 12,000 g for 10 minutes.
9. Transfer supernatant to a new 1.5 ml microcentrifuge. Discard the previous tube.
10. Repeat steps 7-9.

Precipitation of DNA

11. Add 680 µl cold isopropanol (volume ratio 1:1). Mix by inverting gently 50 times. Centrifuge at 8000 g for 5 minutes.
12. Carefully discard the supernatant, avoiding contact with the pellet. Drain the tube by placing on sterile blue roll, taking care that the pellet remains in the tube.

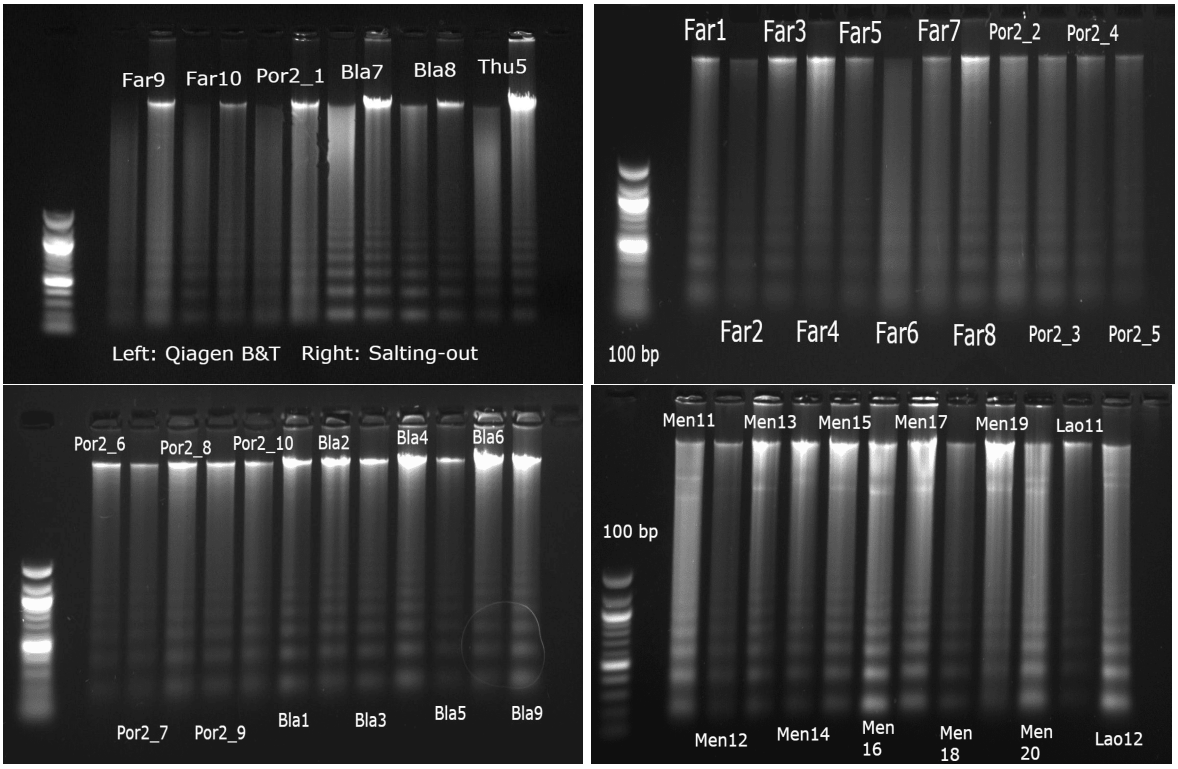
Washing of DNA

13. Add 400 µl 70 % ethanol. Invert the tube several times to wash the DNA pellet. Centrifuge at 8000 g for 1 minute.
14. Carefully discard the supernatant, avoiding contact with the pellet. If a lot of supernatant remains, pulse centrifuge the tubes and discard the supernatant using a smaller pipette, again avoiding contact with the pellet.
15. Allow to air dry to 10-20 minutes. Non-contaminated pellets will turn more transparent as they dry. Avoid over-drying the DNA pellet, as the DNA will be difficult to dissolve.

Rehydration of DNA

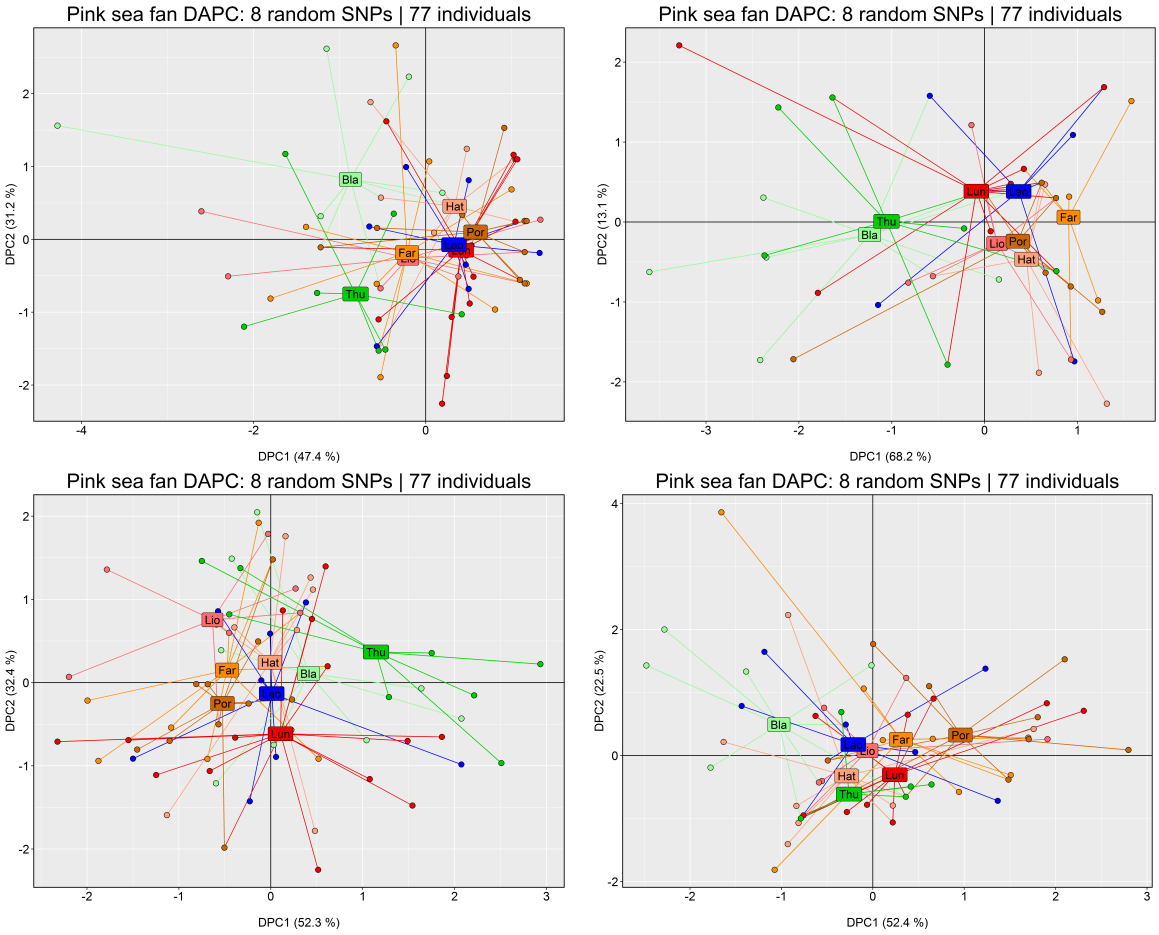
16. Re-suspend dried pellets with 100 µl nuclease-free water. Invert tube to mix and spin down using centrifuge.
17. Incubate at room temperature for 30 minutes or incubate in the fridge overnight.
18. Briefly pulse centrifuge tubes and store at -20^{o}C.

A5: Pink sea fan gel images



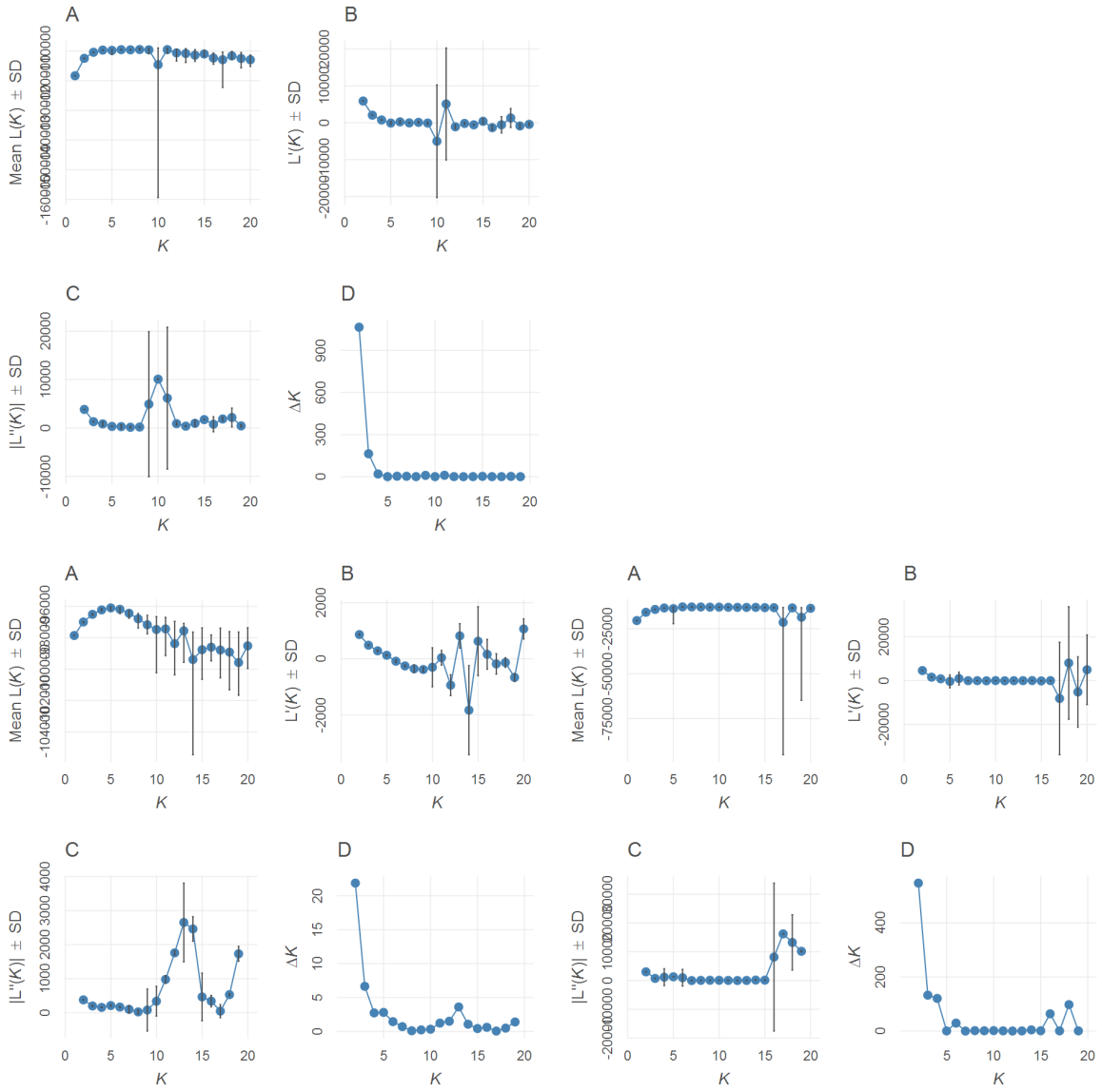
Pink sea fan DNA run on a 1 % agarose gel and stained with ethidium bromide. The top-left image compares the performance of the Qiagen Blood and Tissue kit and the salting-out protocol; the other three images show DNA samples extracted using the salting-out protocol.

A6: Pink sea fan DAPC using eight random SNPs



Pink sea fan discriminant analysis of principle components (DAPC) using eight randomly selected SNPs.

A7: European lobster SNP dataset: interpreting *K*



Interpreting *K* for the European lobster SNP dataset using (A) mean L(*K*), (B) L′(*K*), (C) L″(*K*), (D) delta *K*. Each plot represents analysis with all 86 SNPs (top), analysis with 71 putatively neutral SNPs (bottom-left), and analysis with 15 outlier SNPs (bottom-right).

Meta-analysis of northeast Atlantic marine taxa shows contrasting phylogeographic patterns following post-LGM expansions

Tom L. Jenkins[1], Rita Castilho[2] and Jamie R. Stevens[1]

[1] Department of Biosciences, University of Exeter, Exeter, United Kingdom
[2] Center for Marine Sciences, Campus de Gambelas, University of Algarve, Faro, Portugal

ABSTRACT

Background. Comparative phylogeography enables the study of historical and evolutionary processes that have contributed to shaping patterns of contemporary genetic diversity across co-distributed species. In this study, we explored genetic structure and historical demography in a range of coastal marine species across the northeast Atlantic to assess whether there are commonalities in phylogeographic patterns across taxa and to evaluate whether the timings of population expansions were linked to the Last Glacial Maximum (LGM).

Methods. A literature search was conducted using Web of Science. Search terms were chosen to maximise the inclusion of articles reporting on population structure and phylogeography from the northeast Atlantic; titles and abstracts were screened to identify suitable articles within the scope of this study. Given the proven utility of mtDNA in comparative phylogeography and the availability of these data in the public domain, a meta-analysis was conducted using published mtDNA gene sequences. A standardised methodology was implemented to ensure that the genealogy and demographic history of all mtDNA datasets were reanalysed in a consistent and directly comparable manner.

Results. Mitochondrial DNA datasets were built for 21 species. The meta-analysis revealed significant population differentiation in 16 species and four main types of haplotype network were found, with haplotypes in some species unique to specific geographical locations. A signal of rapid expansion was detected in 16 species, whereas five species showed evidence of a stable population size. Corrected mutation rates indicated that the majority of expansions were estimated to have occurred after the earliest estimate for the LGM (~26.5 Kyr), while few expansions were estimated to have pre-dated the LGM.

Conclusion. This study suggests that post-LGM expansion appeared to be common in a range of marine taxa, supporting the concept of rapid expansions after the LGM as the ice sheets started to retreat. However, despite the commonality of expansion patterns in many of these taxa, phylogeographic patterns appear to differ in the species included in this study. This suggests that species-specific evolutionary processes, as well as historical events, have likely influenced the distribution of genetic diversity of marine taxa in the northeast Atlantic.

Submitted 30 May 2018
Accepted 30 August 2018
Published 28 September 2018

Corresponding author
Jamie R. Stevens, j.r.stevens@ex.ac.uk

Academic editor
Mark Costello

Additional Information and Declarations can be found on page 17

DOI **10.7717/peerj.5684**

How to cite this article Jenkins et al. (2018), Meta-analysis of northeast Atlantic marine taxa shows contrasting phylogeographic patterns following post-LGM expansions. **PeerJ 6:e5684; DOI 10.7717/peerj.5684**

Subjects Biogeography, Genetics, Marine Biology
Keywords Comparative phylogeography, Historical demography, Last Glacial Maximum, mtDNA, Northeast Atlantic, Population expansion

INTRODUCTION

Comparative phylogeographic studies present opportunities to explore how historical events may have helped shape patterns of genetic structure amongst co-distributed species (*Avise et al., 1987*; *Avise, 2009*; *Hickerson et al., 2010*). Patterns of concordant phylogeographical structure across multiple taxa are particularly informative because, while some patterns of spatial genetic structure may be caused by species-specific evolutionary processes, patterns common across multiple taxa may suggest similar evolutionary histories, such as common barriers to gene flow (*Avise, 2009*; *Hickerson et al., 2010*). These findings can be important for conservation because of the potential to modify management actions in the light of the differing phylogeography of multiple species across the same geographical area (*Pelc, Warner & Gaines, 2009*; *Toonen et al., 2011*; *Heyden et al., 2014*; *Liggins et al., 2016*). In marine biology, such comparative studies have made important contributions to our understanding of how historical events, such as the Pleistocene glaciations, have helped shape the spatial patterns of contemporary genetic diversity of marine taxa (*Patarnello, Volckaert & Castilho, 2007*; *Maggs et al., 2008*; *Marko et al., 2010*; *Ni et al., 2014*).

The Pleistocene epoch was characterised by recurrent glaciations and intensive fluctuations in climate that periodically influenced the spatial distributions of plants and animals (*Hewitt, 1999*; *Hofreiter & Stewart, 2009*). The most recent glacial period began approximately 115 Ka and nearly all ice sheets were at their maximum (Last Glacial Maximum, LGM) between 26.5–19 Ka (*Clark et al., 2009*). The advances of the Northern Hemisphere ice sheets led to significant changes in temperature and sea levels (*Lambeck & Chappell, 2001*). This must have had profound implications for habitat availability and the population persistence of coastal species—large parts of species' ranges would have been reduced, while other species may have survived in glacial refugia (*Maggs et al., 2008*; *Provan & Bennett, 2008*). As the ice retreated and the sea level rose, a number of individuals from refugial populations may have dispersed and recolonised areas unavailable during the glaciation (*Hewitt, 2000*). Changes in latitudinal ranges and population sizes can have distinct effects on the genetic architecture of a species due to the competing processes of mutation, drift and selection; moreover, the deep molecular divergence reported in taxa associated with several known European refugia suggests repeated expansion and contraction of conspecific populations were common throughout the Pleistocene (*Hewitt, 2004*).

In the northeast Atlantic, the ice sheets extended as far south as Britain and Ireland, leaving an ice-free zone in mid-southern England, with possibly a small area in southwest Ireland free of ice (*Chiverrell & Thomas, 2010*). However, the predicted extent of ice coverage across southern Ireland and the Celtic Sea differs among studies (e.g., *Taberlet et al., 1998*; *Hughes et al., 2016*). The advance of the ice sheets led to a drastic drop in sea levels in the English Channel, resulting in the complete emersion of the channel between

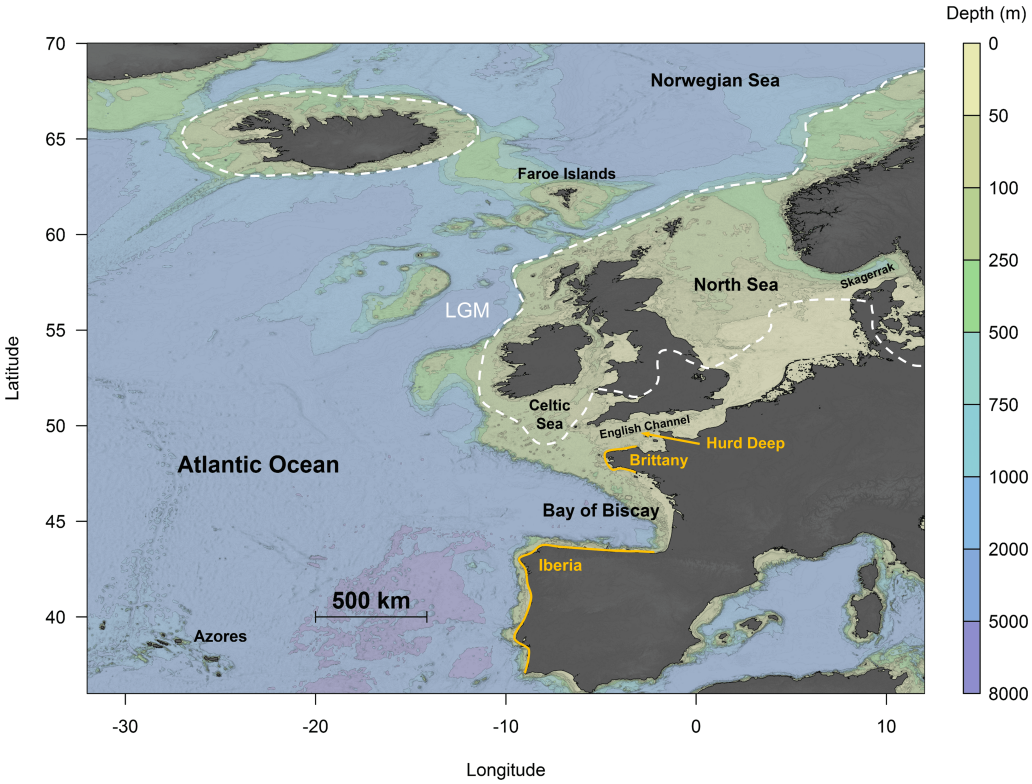


Figure 1 Topographical map of the northeast Atlantic Ocean. The white dotted lines represent the maximum extent of ice cover during the Last Glacial Maximum (LGM) (redrawn from *Hughes et al., 2016*). Orange lines indicate putative refugia: Hurd Deep, Brittany and Iberia.

Full-size DOI: 10.7717/peerj.5684/fig-1

England and France, except for a palaeo-river that extended across the continental margin (*Ménot et al., 2006*). This suggests that extant coastal communities inhabiting these areas are likely recolonisers originating from glacial refugia. It has been suggested that Hurd Deep, a trench in the English Channel (Fig. 1), might have persisted as a marine lake during the LGM, thereby acting as a potential glacial refugium (*Provan, Wattier & Maggs, 2005*; *Hoarau et al., 2007*). Other areas further south, including Brittany (*Coyer et al., 2003*) and the Iberian Peninsula (*Hoarau et al., 2007*; *Neiva et al., 2012*) (Fig. 1), have also been postulated to act as refugia during the LGM. This was supported by high levels of genetic diversity found at these areas in the species studied, a key signature indicative of glacial refugia (*Provan & Bennett, 2008*).

Studies of single-species phylogeography across the northeast Atlantic are common; yet, because of the differences in molecular methodologies and analytical approaches, it can be difficult to compare results reliably. By applying a consistent methodology across all studies, this standardises the analysis (*Harrison, 2011*), enabling patterns of phylogeography to be explored and compared within and across taxa. Two comparative meta-analyses in the Atlantic Ocean have been published to-date: the first explored the feasibility of distinguishing genetic signatures of periglacial refugia from southern refugia

in eight benthic marine species (*Maggs et al., 2008*), and the second looked for concordance among phylogeographical breaks around the southeast coast of the United States of America (*Pelc, Warner & Gaines, 2009*). Systematic meta-analyses across diverse taxa in other seas and oceans have proved useful for exploring broad patterns of phylogeography (e.g., *Patarnello, Volckaert & Castilho, 2007*; *Kelly & Palumbi, 2010*; *Marko et al., 2010*; *Ni et al., 2014*); for example, one study of rocky-shore taxa from the northeastern Pacific found that 36% of species showed evidence of population expansions associated with the LGM, while 50% exhibited demographic patterns consistent with stable effective population sizes (*Marko et al., 2010*). However, such a study for marine taxa across the northeast Atlantic has yet to be undertaken.

In this study, we reanalyse available mitochondrial (mt)DNA data to compare the phylogeography of coastal benthic and demersal organisms across the northeast Atlantic (Fig. 1), an area characterised by complex oceanography and historical biogeographical events, such as the Pleistocene glaciations. Specifically, our aims were: (i) to identify commonalties (or otherwise) in contemporary genetic structure; (ii) to re-examine historical demography to test for signatures of population expansions; and (iii) to estimate the timings of any expansions detected. We discuss our findings in the context of the Pleistocene glaciations, asking in particular whether the LGM affected the phylogeography of marine taxa concordantly or discordantly.

MATERIAL AND METHODS

Literature search

To compare the phylogeography of benthic and demersal organisms across the northeast Atlantic, we undertook a meta-analysis of molecular phylogeographic studies. A literature search was conducted by TLJ and JRS using Web of Science (Thomson Reuters) in February 2015. Search terms were chosen to maximise the inclusion of articles reporting on population structure and phylogeography from the northeast Atlantic. The following sets of Boolean search terms were submitted to the Advanced Search Tool: (1) gene flow OR population structure OR genetic diversity OR phylogeograph*; (2) marine OR intertidal OR subtidal OR estuar*; and (3) Atlantic. Titles and abstracts were screened by TLJ and JRS to identify suitable articles within the scope of this study and only articles that matched the following criteria were retained: (a) organisms were fully marine or estuarine throughout their life history (diadromous species were excluded); (b) studies of temporal changes, hybridisation or introgression from closely related species were omitted; (c) the study included at least three sampling sites from within the northeast Atlantic (Fig. 1—sites outside of this area were not considered); (d) datasets contained a minimum of five individuals per site and a total sample size of at least 50; and (e) the study included latitude and longitude of the sampling sites or a detailed description or map which provided sufficient detail to determine the geographical location of sample origins. The studies were reviewed independently by TLJ and JRS and there were no disputes regarding inclusion or rejection that needed adjudication. Given the proven utility of mtDNA in comparative phylogeography (e.g., *Patarnello, Volckaert & Castilho, 2007*; *Ni et al., 2014*)

and the availability of these data in the public domain, a meta-analysis was conducted using published mtDNA gene sequences.

Data reanalysis

A standardised methodology was implemented to ensure that all mtDNA datasets were reanalysed in a consistent and directly comparable manner. Data analyses in the original studies were far from consistent, particularly with respect to the analysis of haplotype networks and historical demography. The majority of studies reported information about population structure, however, in several instances the studies included additional samples outside of the northeast Atlantic in their analysis. Therefore, standardised tests of population structure were undertaken *de novo* for each species. Sites that were genetically homogeneous (as described by the original authors) and which were spatially close or situated in the same geographical region were combined in some datasets. This ensured that phylogeography within and across seas was examined in this meta-analysis. Population differentiation was examined using global values of Jost's D (*Jost, 2008*) and F_{ST} (*Weir & Cockerham, 1984*) using the *fastDivPart* function from the R package diveRsity (*Keenan et al., 2013*; *R Core Team, 2016*) and significance was assessed using 10,000 permutation replicates.

To examine the genealogical relationships within species, haplotype networks were constructed using the *haploNet* function from the R package pegas (*Paradis, 2010*). Tajima's D (*Tajima, 1989*), Fu's F_S (*Fu, 1997*) and Ramos-Onsins' R_2 (*Ramos-Onsins & Rozas, 2002*) neutrality tests were performed in DnaSP v5.10 (*Librado & Rozas, 2009*) to determine whether each species carried a signal that deviated from neutrality (significance was assessed using 10,000 bootstrap replicates). Mismatch analyses (frequency of pairwise nucleotide-site differences between sequences) were carried out using the population growth-decline model in DnaSP to further examine the demographic history, and Harpending's raggedness index (r) (*Harpending, 1994*) was used to evaluate the fit of the observed distribution to the growth-decline model (10,000 bootstrap replicates). A non-significant index suggests that the observed data have a relatively good fit to the growth-decline model. In contrast, a significant index is indicative of a stable population which is typically thought to show a 'ragged', multi-modal mismatch (*Harpending, 1994*).

The equation $t = \tau/(2\mu k)$ was used to estimate the timing of a population expansion (t), where τ is the date of the expansion measured in units of mutational time (Tau –estimated using DnaSP), $\mu$is the mutation rate per site per year and k is the sequence length. In addition, Bayesian Skyline Plots (BSPs) were run using BEAST2 v2.5.0 (*Drummond et al., 2005*; *Bouckaert et al., 2014*). BEAST2 uses a Markov chain Monte Carlo (MCMC) sampling procedure to estimate effective population size (N_e) through time based on the temporal distribution of coalescences in gene genealogies. For each dataset, the substitution model was selected using bModelTest (*Barido-Sottani et al., 2018*), which uses reversible jump MCMC that allows the Markov chain to jump between states representing different possible substitution models. A strict clock and a coalescent Bayesian Skyline prior was implemented. Each run consisted of 100 million steps with a burn-in of one million and

parameters were sampled every 10,000 steps. Chain convergence and BSPs were analysed with Tracer v1.7.1 (*Rambaut et al., 2018*).

Recent studies have shown that the use of mutation rates derived from ancient calibration dates or from phylogenetic analyses may not be appropriate for studies at the population level (*Ho et al., 2008*; *Ho et al., 2011*). In this study, therefore, mutation rates were chosen based on the most recent calibration date available for the closest taxonomic relative (Table S1). In published studies where a mutation rate was not specified, the genetic distance provided by the study was divided by the date of the calibration event (in Myr) to obtain a % mutation rate per Myr. For cases where only calibration dates older than 5 Myr were available for the species and gene of interest, a three-fold correction in mutation rate was applied to the original rate to control for the potential time-dependency of molecular rates. This adjustment was implemented because rates have been found to vary by three to six-fold for several marine species when calibration dates younger than 5 Myr vs. older dates have been tested (*Crandall et al., 2012*; *Laakkonen, Strelkov & Väinölä, 2015*). A range of mutation rates based on the rates reported by previous studies were used to calculate a minimum, maximum and average time estimate since a population expansion.

RESULTS

Literature search

The initial search using Boolean terms identified 1,120 articles, which was reduced to 56 articles after the titles and abstracts were examined and the search criteria were applied (Fig. S1). The final database for the meta-analysis consisted of mtDNA gene sequence data from 21 studies (Table 1); some studies from the previous step were not included due to the use of RFLPs in mtDNA or because some mtDNA datasets were not publicly available. The final database spanned several taxonomic groups, with fishes, molluscs and crustaceans accounting for the majority of species (81%). The most common mitochondrial gene across all studies was cytochrome oxidase I (COI), followed by cytochrome *b* (Cyt *b*), the control region (CR) and the intergenic spacer region (IGS). COI was the most commonly used gene for invertebrate studies, IGS for macroalgae, and studies of fish used either the CR or the Cyt *b* gene.

Genetic structure

Sixteen species showed significant global Jost's *D* and F_{ST} values, indicative of population differentiation (Table 2), while the remaining five species showed little evidence of population differentiation. Across the 21 datasets, four different types of haplotype network were putatively identified based on the structure of the networks (Fig. 2) (all haplotype networks are presented in Fig. S2):

(i) A 'Star' network (Fig. 2A), in which a single, widespread haplotype is typically positioned at the centre of the network and is thought to be the ancestral haplotype. Additional haplotypes are linked to this dominant haplotype by a single (or a few) mutational step(s), suggesting these haplotypes are the product of recent mutation events. Eight species showed this type of relationship (*Celleporella hyalina*, *Conger conger*, *Nassarius nitidus*, *Nassarius reticulatus*, *Palinurus elephas*, *Pelvetia canaliculata*, *Pomatoschistus*

Table 1 List of the papers used in the meta-analysis and a summary of the information extracted from each study.

Taxon species	MtDNA gene	No. sites; *N*	Sampling site distribution	Larval development	No. of lineages	Reference
Crustacean						
Carcinus maenas	COI	13; 200	SW Spain to Norway	PLD, long	1	*Roman & Palumbi (2004)*
Maja brachydactyla	COI	13; 291	SW Spain to W Ireland	PLD, 2–3 wk	1	*Sotelo et al. (2008)*
Neomysis integer	COI	9; 379	SW Spain to E Scotland	No PLD, brooder	1	*Remerie et al. (2009)*
Palinurus elephas	COI	6; 119	S Portugal to W Scotland	PLD, up to 1 yr	1	*Palero et al. (2008)*
Fish						
Conger conger	CR	4; 232	Azores to Ireland	Leptocephalus, up to 2 yr	1	*Correia et al. (2012)*
Dicentrarchus labrax	CR	9; 93	Bay of Biscay to Norway	PLD, 8–12 wk	1	*Coscia & Mariani (2011)*
Labrus bergylta	CR	7; 279	W Ireland to Norway	PLD, 37–49 d	1	*D'Arcy, Mirimin & FitzGerald (2013)*
Pomatoschistus microps	Cyt *b*	10; 232	Bay of Biscay to Norway	PLD, 6–9 wk	1	*Gysels et al. (2004)*
Pomatoschistus minutus	Cyt *b*	8; 165	S Portugal to Norway	PLD, unknown	1	*Larmuseau et al. (2009)*
Raja clavata	Cyt *b*	9; 315	Azores to North Sea	No PLD, oviparous	1	*Chevolot et al. (2006)*
Solea solea	Cyt *b*	10; 645	Bay of Biscay to Skagerrak	PLD, up to 3 wk	1	*Cuveliers et al. (2012)*
Symphodus melops	CR	10; 263	S Portugal to Skagerrak	PLD, 14–25 d	1	*Robalo et al. (2012)*
Macroalgae						
Pelvetia canaliculata	IGS	15; 429	Portugal to Norway	No PLD, external fertilisation	1	*Neiva et al. (2014)*
Mollusc						
Cerastoderma edule	COI	12; 300	Portugal to Norway	PLD, up to 4 wk	1	*Krakau et al. (2012)*
Macoma balthica	COI	15; 339	Bay of Biscay to North Sea	PLD, 2–5 wk	2	*Becquet et al. (2012)*
Modiolus modiolus	COI	4; 73	Irish Sea to Norway	PLD, up to 24 wk	2	*Halanych et al. (2013)*
Nassarius nitidus	COI	3; 62	NW Spain to Sweden	PLD, 4–8 wk	1	*Couceiro et al. (2012)*
Nassarius reticulatus	COI	6; 156	S Portugal to UK	PLD, 4–8 wk	1	*Couceiro et al. (2007)*
Polychaete						
Owenia fusiformis	COI	11; 283	Portugal to North Sea	PLD, up to 28 d	3	*Jolly et al. (2005)*
Pectinaria koreni	COI	10; 289	Portugal to North Sea	PLD, up to 15 d	2	*Jolly et al. (2006)*
Bryozoan						
Celleporella hyalina	COI	9; 63	NW Spain to Iceland	PLD, 1–4 h	1	*Gómez et al. (2007)*

Notes.
MtDNA, mitochondrial DNA; No. of sites, number of sampling sites; *N*, total number of sequenced individuals; PLD, pelagic larval duration.

Table 2 **Summary statistics for each species. Population differentiation and demographic statistics are shown. In all statistical tests, significance was assessed using 10,000 permutations or bootstraps replicates.**

Species	Population differentiation		Demography				
	Jost's *D*	F_{ST}	Tajima's *D*	F_S	R_2	*r*	Expansion
Crustacean							
Carcinus maenas	0.584***	0.157***	−1.73*	−40.36***	0.034*	0.018	Yes
Maja brachydactyla	0.298***	0.045***	−1.86**	−33.72***	0.028*	0.030	Yes
Neomysis integer	0.956***	0.554***	0.14	−0.954	0.024	0.086	No
Palinurus elephas	0.023	0.000	−2.31***	−30.19***	0.019*	0.094	Yes
Fish							
Conger conger	0.124	0.000	−2.58***	−211.1***	0.012***	0.031	Yes
Dicentrarchus labrax	0.540*	0.031*	−1.88**	−21.52***	0.047*	0.011	Yes
Labrus bergylta	0.672***	0.135***	−0.53	−49.35***	0.074	0.024	Yes
Pomatoschistus microps	0.391***	0.385***	−1.39	−17.90***	0.044	0.215	Yes
Pomatoschistus minutus	0.652***	0.100***	−1.96**	−90.56***	0.034*	0.015	Yes
Raja clavata	0.375***	0.330***	−0.09	−2.340	0.076	0.309	No
Solea solea	0.049	0.002	−2.02***	−131.9***	0.021**	0.221	Yes
Symphodus melops	0.578***	0.349***	−1.70*	−50.52***	0.032*	0.086	Yes
Macroalgae							
Pelvetia canaliculata	0.689***	0.482***	−1.53*	−19.02***	0.036	0.043	Yes
Mollusc							
Cerastoderma edule	0.662***	0.304***	−2.24***	−34.47***	0.019**	0.033	Yes
Macoma balthica	0.702***	0.470***	–	–	–	–	–
lineage 1	–	–	−0.80	−3.773	0.053	0.241	No
lineage 2	–	–	−0.99	−1.110	0.089	0.173	No
Modiolus modiolus[a]	0.083	<0.001	−1.79*	−11.91***	0.045*	0.156	Yes
Nassarius nitidus	0.222***	0.302***	−1.49*	0.028	0.049*	0.446	No
Nassarius reticulatus	0.047	0.000	−2.51***	−48.33***	0.016**	0.080	Yes
Polychaete							
Owenia fusiformis	0.788***	0.055***	–	–	–	–	–
lineage 1	–	–	−2.34***	−114.8***	0.024**	0.020	Yes
lineage 2	–	–	−2.06**	−55.00***	0.030**	0.008**	Yes
lineage 3	–	–	−1.26	−3.934**	0.084	0.080	Yes
Pectinaria koreni	0.596***	0.112***	–	–	–	–	–
lineage 1	–	–	−1.99**	−76.48***	0.027**	0.021	Yes
lineage 2	–	–	−2.63***	−54.02***	0.018***	0.029*	Yes
Bryozoan							
Celleporella hyalina	0.513***	0.488***	−1.35	−0.554	0.063	0.061	No

Notes.

*<0.05.

**<0.01.

***<0.001.

F_s, Fu's F_s; R_2, Ramos-Onsins' R_2; *r*, Harpending's raggedness index.

[a]Only statistics for lineage 1 are shown.

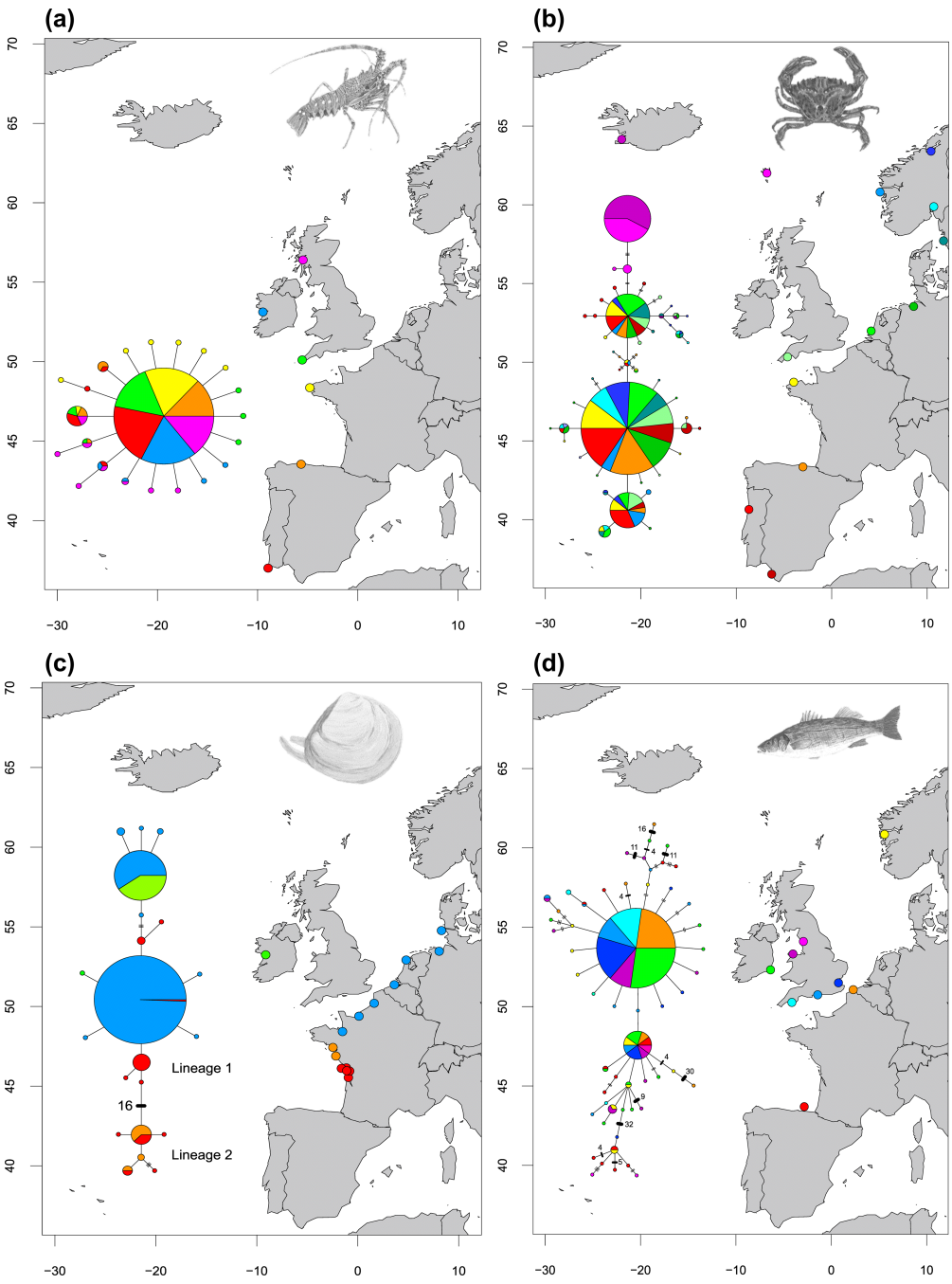


Figure 2 Haplotype networks showing four different network structures. Haplotype networks showing (A) 'star' (*Palinurus elephas*), (B) 'complex star' (*Carcinus maenas*), (C) 'reciprocally monophyletic' (*Macoma balthica*) and (D) 'complex mutational' (*Dicentrarchus labrax*) structures. Each circle represents a unique haplotype and the sizes of the circles are proportional to the haplotype frequencies for each network but are not comparable across studies. Each line represents one mutation step and two or more steps are indicated by bars or numbers. Colours inside the circles correspond to sites which have individuals represented in that particular haplotype. Species illustrations by Guy Freeman.

Full-size DOI: 10.7717/peerj.5684/fig-2

microps and *Raja clavata*). In one case, the dominant haplotype had far fewer connections than a low-frequency haplotype in the network, making it difficult to distinguish the centre of the network with confidence (*Pomatoschistus microps*);

(ii) A 'Complex star' network (Fig. 2B), in which there are multiple high-frequency haplotypes and connections. Six species showed this type of relationship (*Carcinus maenus*, *Cerastoderma edule*, *Maja brachydactyla*, *Pomatoschistus minutus*, *Solea solea, Symphodus melops*);

(iii) A 'Reciprocally monophyletic' network (Fig. 2C), in which more than one lineage is apparent and each lineage is linked by a long branch associated with numerous mutations. Four species showed this type of relationship (*Macoma balthica*, *Modiolus modiolus*, *Owenia fusiformis* and *Pectinaria koreni*);

(iv) A 'Complex mutational' network (Fig. 2D), in which some branches were separated by a very large number of mutations, while other branches had contrarily one or two mutations. Three species showed this type of relationship (*Dicentrarchus labrax*, *Labrus bergylta* and *Neomysis integer*). In most cases, a dominant haplotype was present and was presumed to be the ancestral form. However, *Neomysis integer* presented an unusual network in which a distinct ancestral haplotype was not apparent and the centre of the haplotype network was not readily distinguishable.

Historical demography

Historical demography was inferred for each species based on the observed mismatch distribution, neutrality tests and the raggedness index (Table 2). Four main types of mismatch distributions were observed: unimodal, skewed unimodal, multimodal and bimodal (Fig. 3) (all mismatch distributions are presented in Fig. S3). Unimodal is associated with a sudden population expansion (e.g., *Maja brachydactyla*; Fig. 3A), and skewed unimodal is generally associated with a recent expansion or bottleneck (e.g., *Nassarius reticulatus*; Fig. 3B). Multimodal (e.g., *Labrus bergylta*; Fig. 3C) and bimodal (e.g., *Macoma balthica*; Fig. 3D) are usually associated with constant population size. However, previous research has suggested that bimodal peaks may indicate the presence of two distinct lineages (e.g., *Alvarado-Bremer et al., 2005*), which would potentially violate the assumptions of coalescent theory if analysed as one 'genetic' population. In this case, the first peak would represent intra-clade pairwise differences, whereas the second peak would likely represent more ancient inter-clade pairwise differences (Fig. 3D). For each instance of bimodality, the haplotype network was inspected for evidence of two or more lineages. The networks indicated that more than one distinct lineage was evident for all bimodal mismatches (*Macoma balthica*, *Modiolus modiolus*, *Owenia fusiformis* and *Pectinaria koreni*) and, therefore, mismatch analysis and neutrality tests were carried out on each lineage separately. These analyses were not conducted for lineage 2 of *Modiolus modiolus* due to the small number of individuals ($N = 3$) comprising this lineage.

Neutrality statistics for testing the drift–mutation equilibrium (Tajima's D, F_S and R_2) were found to be contrasting between species (Table 2). These tests tended to be significant for species that showed a star-shaped network and for which the mismatch graph was unimodal or skewed unimodal. This supported evidence that a signal of rapid population

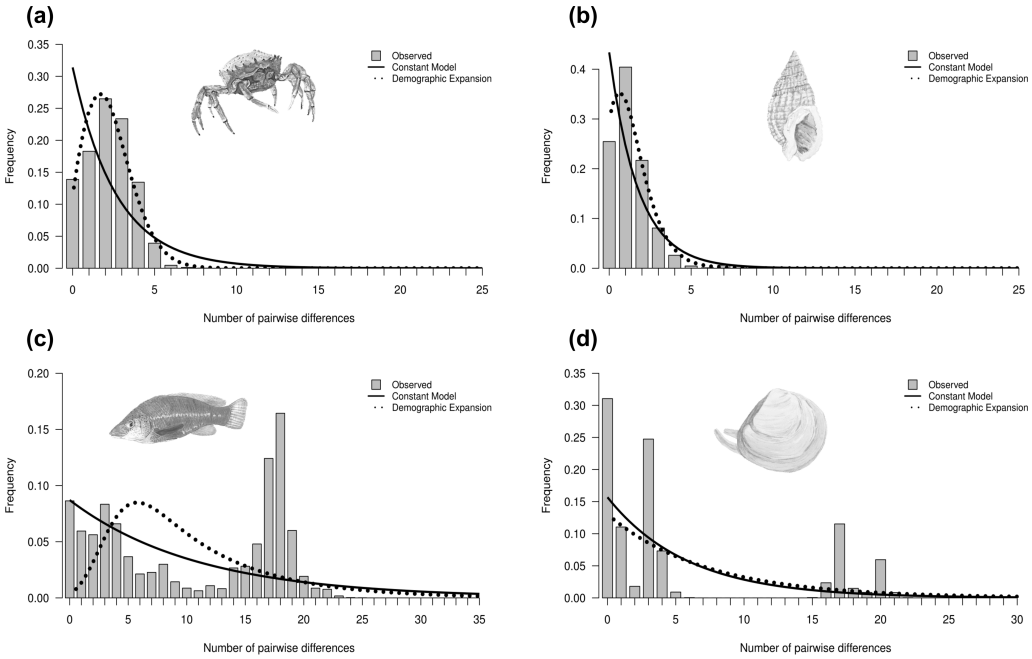


Figure 3 Mismatch distributions showing four different distributions. Mismatch distributions showing (A) unimodal (*Maja brachydactyla*), (B) skewed unimodal (*Nassarius reticulatus*), (C) multimodal (*Labrus bergylta*) and (D) bimodal (*Macoma balthica*). Unimodal and skewed unimodal distributions are generally associated with a sudden expansion and a recent sudden expansion, respectively. Multimodal and bimodal are thought to be associated with a constant population size (but see text). Bars represent the frequency of pairwise nucleotide differences between individuals. Curves correspond to the expected distribution fitted to the data under a model of constant population size (solid line) or demographic expansion (dotted line). Species illustrations by Guy Freeman.

Full-size DOI: 10.7717/peerj.5684/fig-3

expansion was detected; however, a selective sweep can also produce the same genetic signal. Harpending's *r* suggested that two datasets departed from a model of demographic expansion (Table 2), but inspection of the mismatch graphs and neutrality tests indicated there was strong evidence to support a rapid population expansion (or selective sweep) in both datasets. No signatures of rapid population expansion were detected in five species (*Celleporella hyalina*, *Macoma balthica*, *Nassarius nitidus*, *Neomysis integer* and *Raja clavata*), suggesting a stable constant population size.

For the remaining 19 datasets (16 species, 19 including lineages), a historic population expansion was assumed and the timing of the expansion was estimated (Fig. 4). All expansions were found to take place during the Pleistocene or the Holocene epoch. Estimated timings for 17 datasets were after or overlapped the earliest estimate for the LGM (~26.5 Ka). Expansion estimates for one fish (*Labrus bergylta*) and one lineage of the polychaete *Owenia fusiformis* pre-dated the LGM but were still positioned during the last glacial period. Bayesian Skyline Plots (Fig. 5) were generally consistent with the results from the mismatch analyses. Among the 17 datasets for which from the mismatch analyses expansion times were estimated to have occurred after the LGM, a rise in N_e post-LGM was apparent in 15 of these datasets, but the strength of the increase varied across datasets.

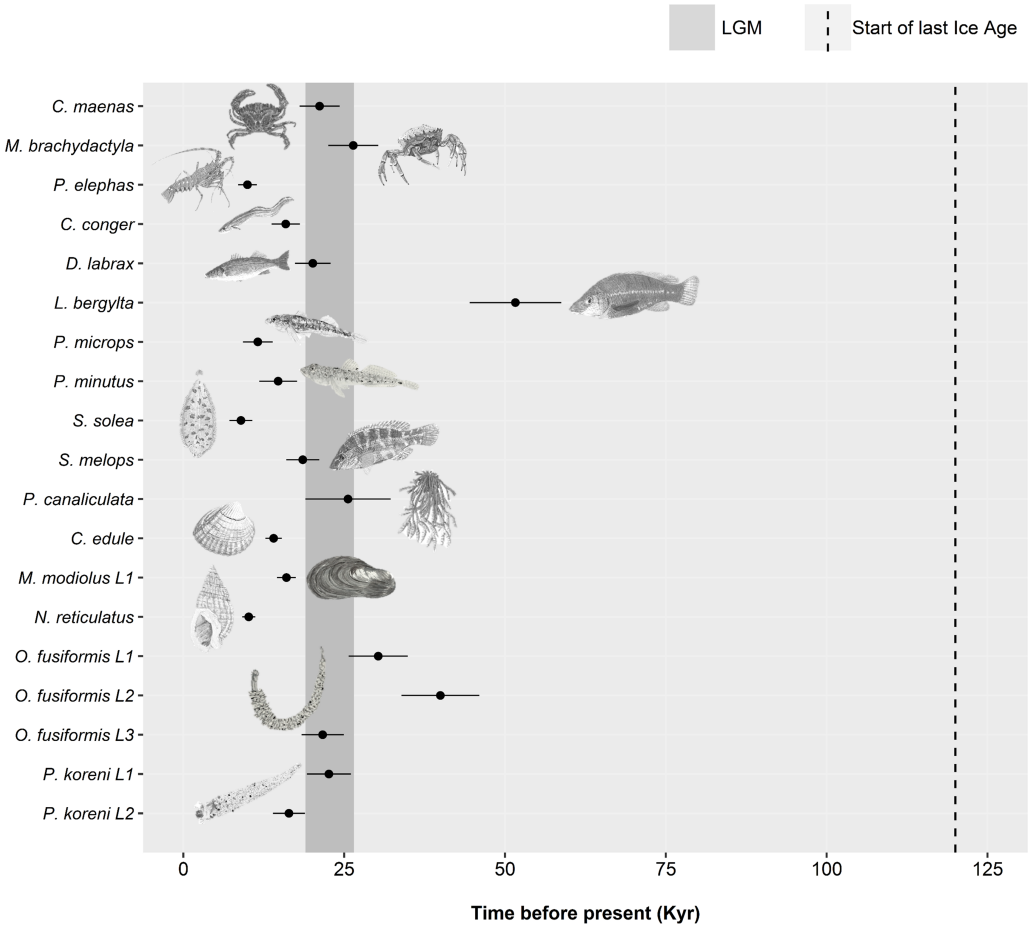


Figure 4 Estimated dates of expansion for species or lineages (L) in which the demographic expansion hypothesis was not rejected. A minimum and maximum time since expansion is plotted as horizontal bars for some datasets, estimated from a minimum and maximum mutation rate (Table S1). The beginning of the last glacial period (dotted line) and the estimated time-frame of the Last Glacial Maximum (grey shaded area) are displayed. Species are organised by taxa: crustaceans, *Carcinus maenas* –*Palinurus elephas*); fish, *Conger conger* –*Symphodus melops*; macroalgae, *Pelvetia canaliculata*; molluscs, *Cerastoderma edule* –*Nassarius reticulatus*; polychaetes, *Owenia fusiformis* –*Pectinaria koreni*. Species illustrations by Guy Freeman.

Full-size DOI: 10.7717/peerj.5684/fig-4

In comparison to the mismatch analysis, the BSP for *L. bergylta* (Fig. 5F) and *O. fusiformis* lineage 2 (Fig. 5P) indicated a population expansion after the earliest estimate for the LGM as opposed to pre-dating the LGM. In addition, although the mismatch analyses inferred a post-LGM expansion for *M. modiolus* lineage 1 (Fig. 5M) and *O. fusiformis* lineage 3 (Fig. 5Q), BSPs generally suggested N_e was constant after the LGM.

DISCUSSION

The results of this study show a range of contemporary genetic patterns across the coastal marine taxa analysed in the northeast Atlantic. In general, genealogical patterns were

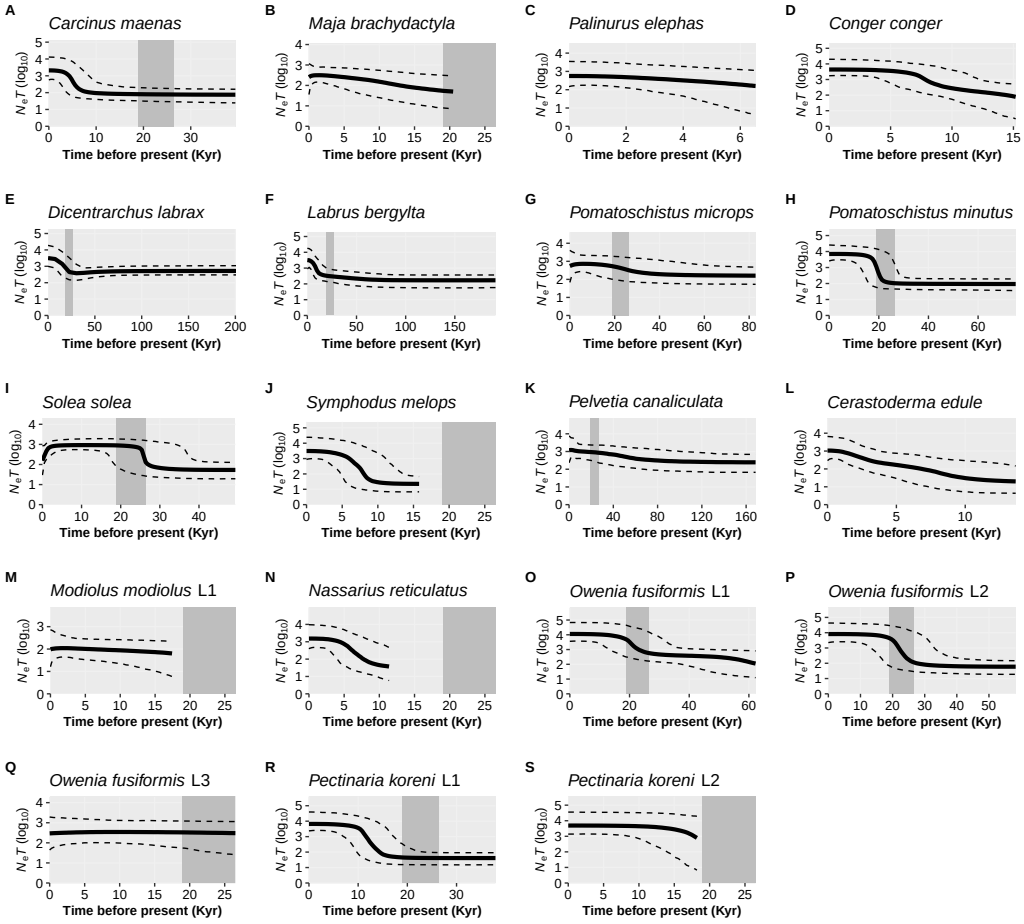


Figure 5 Bayesian Skyline Plots for species or lineages (L) in which the demographic expansion hypothesis was not rejected. Solid black lines show the medium effective population size over time (N_e = effective population size and T = generation time); dashed black lines represent the 95% confidence intervals. The estimated time-frame of the Last Glacial Maximum is denoted by the area shaded dark grey. Species are organised by taxa: crustaceans, *Carcinus maenas* (A), *Maja brachydactyla* (B), *Palinurus elephas* (C); fish, *Conger conger* (D), *Dicentrarchus labrax* (E), *Labrus bergylta* (F), *Pomatoschistus microps* (G), *P. minutus* (H), *Solea solea* (I), *Symphodus melops* (J); macroalgae, *Pelvetia canaliculata* (K); molluscs, *Cerastoderma edule* (L), *Modiolus modiolus* Lineage 1 (M), *Nassarius reticulatus* (N); polychaetes, *Owenia fusiformis* Lineage 1 (O), *O. fusiformis* Lineage 2 (P), O. *fusiformis* Lineage 3 (Q), *Pectinaria koreni* Lineage 1 (R), *P. koreni* Lineage 2 (S).

Full-size DOI: 10.7717/peerj.5684/fig-5

not uniform within taxonomic groups, though common patterns were observed in both polychaete species, which implies that historical events may have affected these polychaete species similarly. Most species (76%) showed evidence of population structuring, suggestive of restricted contemporary or historical gene flow between the sites studied. Of the species that exhibited no population differentiation, all five species have a pelagic larval phase, with a pelagic larval duration (PLD) ranging from up to three weeks (*S. solea*) to a year or more (*P. elephas* and *C. conger*) (Table 1). However, most of the species that demonstrated significant population differentiation also had a pelagic larval phase, ranging from a

relatively short PLD of 1–4 h (*C. hyalina*) to a relatively long PLD of 8–12 weeks (*D. labrax*) (Table 1). Although speculative, taken altogether, this may suggest that larval development and PLD could be important factors in maintaining gene flow in some, but not all, of these species; however, more evidence is needed to confirm this. Indeed, whether a general correlation exists between PLD and genetic differentiation measures remains unclear because some studies have reported poor correlations between the two (*Weersing & Toonen, 2009*; *Kelly & Palumbi, 2010*; *Riginos et al., 2011*), while other studies have reported the opposite (*Siegel et al., 2003*; *Selkoe & Toonen, 2011*) suggesting that PLD and genetic metrics can indeed reflect scales of dispersal if the sampling design is robust (*Selkoe & Toonen, 2011*). As a result, speculative relationships between PLD and genetic differentiation should be interpreted with caution.

In some of the species studied, certain geographical areas were dominated by a particular haplotype that was rarely or not present in other areas across the sampled range. For example, the green crab *Carcinus maenas* showed highly significant differentiation and distinctive haplotypes in the Faroe Islands and Iceland, a pattern detected by the original authors who subsequently concluded that a deep-water barrier to dispersal in green crabs was the driver of this pattern (*Roman & Palumbi, 2004*). A similar pattern was also observed for two species around western Ireland in the northeast Atlantic. In *Celleporella hyalina* and *Macoma balthica*, distinct haplotypes composed a population around western Ireland; however, unique haplotypes were not apparent in other species analysed in this study with similar sampling coverage (e.g., *Labrus bergylta*, *Palinurus elephas* and *Pelvetia canaliculata*). A discrepancy in genetic structure between species at this spatial scale has also been observed between two temperate octocoral species (*Eunicella verrucosa* and *Alcyonium digitatum*) using microsatellite markers, whereby northwest Ireland samples were found to be genetically isolated from other northeast Atlantic samples in *E. verrucosa*, but not in *A. digitatum* (*Holland, Jenkins & Stevens, 2017*). This suggests that historical or contemporary gene flow between areas in the northeast Atlantic and western Ireland is likely possible, but in some cases the spatial patterns of genetic structure could be influenced by other processes such as strong selection pressures, species-specific life history traits, demographic fluctuations, or range expansions occurring at different times in different species (*Hellberg, 2009*).

Demographic history

Demographic history was variable across species in the northeast Atlantic, as evidenced by both the diverse structuring of the haplotype networks and the observed mismatch distributions within species. The presence of one or more lineages and the complexity of mutational patterns in several networks suggested some species have undergone pronounced changes in their demography and genealogy. Connections with large mutation steps separating some haplotypes are indicative of deep phylogenetic splits in the genealogies and suggests the persistence of old populations in these species. Accumulating new mutations is a relatively slow process and, therefore, sufficient time since coalescence must have elapsed to facilitate these large sequence divergences (*Avise, 2009*).

In the northeast Atlantic, the LGM has often been viewed as a possible explanation for discrepancies in genealogies and for rapid population expansions via recolonisation as glaciers started to retreat from their maximum positions (*Hewitt, 2004*). In this study, we detected rapid expansions in many different taxa, of which the majority were estimated to occur after the LGM. This supports evidence for post-LGM expansions, possibly from periglacial refugia (*Maggs et al., 2008*) or via recolonisation of areas previously affected by the Northern Hemisphere ice sheets. These results are in contrast to the northeast Pacific where regional persistence during the LGM appeared to be common in rocky-shore organisms (*Marko et al., 2010*). The conclusions of several previous studies reanalysed in this meta-analysis also detected rapid expansions (e.g., *Jolly et al., 2006*; *Sotelo et al., 2008*; *Larmuseau et al., 2009*); however, the authors of these studies estimated the dates of these expansions to have occurred pre-LGM. This discrepancy could be due to the differences in mutation rates, whereby the original authors typically used rates derived from ancient calibrations, while in this study we attempted to use more recent calibration dates to correct for the potential time-dependency of molecular rates (*Ho et al., 2011*).

Of course, we acknowledge that the signal of deviation from neutrality we detected may, in some cases, be the result of a selective sweep and not a rapid expansion. This signal could be distinguished by incorporating multi-locus data; nevertheless, given that a variety of species in this study showed similar genealogical patterns consistent with demographic expansion, it seems likely that most of them did indeed experience demographic changes associated with the end of the LGM, rather than selective sweeps. Moreover, distinctive haplotypes were found in several species networks (*Pelvetia canaliculata*, *Pomatoschistus minutus*, *Owenia fusiformis* and *Pectinaria koreni*) to the south of where the Eurasian ice sheet is proposed to have extended during the LGM (Fig. 1). This finding suggests populations of these species may have survived in southern glacial refugia; though, as pointed out by some of the original authors, deep sequence divergences in some species (e.g., *O. fusiformis* and *P. korena*) and the lack of a species-specific molecular clock calibration makes inferences about refugia challenging (*Jolly et al., 2005*; *Jolly et al., 2006*).

It is difficult to suggest an explanation for the two expansions estimated to have pre-dated the LGM (using mismatch analysis), but which fall within the last glacial period. This pattern of pre-LGM expansion has also been reported in a number of previous studies for a variety of marine taxa (e.g., *Hoarau et al., 2007*; *Marko et al., 2010*; *Ni et al., 2014*; *Almada et al., 2017*). One potential explanation for this pattern is that sea level during the last glacial cycle did not decrease uniformly towards the level observed at the LGM, but oscillated rapidly over a period of 60 Ka to 30 Ka (see Fig. 3A in *Lambeck, Esat & Potter, 2002*). Therefore, it may be possible that we are detecting the signature of a population expansion during one of these sudden increases in sea level during the last glacial period. Alternatively, as the BSP analysis inferred a post-LGM expansion for these two datasets, this could be a limitation associated with the mismatch analysis approach, which does not consider genealogy, and may, therefore, produce a less precise estimation. In addition, the sample of genetic diversity for this species may not be representative (*Karl et al., 2012*) or the genetic signal we detected may have been the result of a selective sweep and not a rapid expansion.

The use of single marker mtDNA genealogies and coalescence theory can introduce challenges associated with the interpretation of data and these limitations should be acknowledged (*Karl et al., 2012*). For example, the populations under study may have experienced multiple episodes of growth and decline; however, only the most recent expansion event can be detected using coalescence analysis and, in some cases, these events may not be sufficiently severe to be detected (*Karl et al., 2012*). In addition, coalescent histories can differ amongst loci because they can experience mutation and drift independently. Therefore, analysis of a single gene only gives insight into the coalescent history of that locus, which may not always be representative of population history. Analysis of multiple loci and genomics would help to alleviate these concerns, and would likely provide enhanced resolution for exploring the phylogeography of northeast Atlantic marine fauna.

Although population expansions were detected in a number of species in this study and also in the wider literature, populations of other marine species, including five from this study, have been found to remain stable throughout the LGM. As previously reported, not all coastal marine taxa appear prone to demographic changes during or after ice ages (*Janko et al., 2007*; *Marko et al., 2010*; *Olsen et al., 2010*). It is also important to acknowledge that earlier events in the Pleistocene and more ancient events that pre-date the Pleistocene may have helped shape the contemporary patterns of genealogical structure observed in this study.

CONCLUSION

The findings of this meta-analysis indicate that species in the northeast Atlantic do not show a uniform pattern of phylogeography, but rather a mixture of complex contemporary genealogical structure. Reanalysis of demographic histories indicated that a large proportion of the species included in this study have experienced post-LGM expansions, supporting the general expectation that rapid population expansions occurred after the LGM as the ice sheets started to retreat (*Hewitt, 2000*; *Hewitt, 2004*). This suggests that regional extirpation during the LGM appears to be a common biogeographic history for many northeast Atlantic marine taxa. However, improvements in mutation rate estimates, as well as the incorporation of multi-locus markers and genomics, would likely provide greater accuracy and resolution for overcoming the challenges associated with single mtDNA genealogies, and for improving our understanding of phylogeography in the northeast Atlantic Ocean.

ACKNOWLEDGEMENTS

We thank the many authors who kindly provided or contributed to the original datasets used in the meta-analysis. We also thank SYW Ho (Sydney), CN Roterman (Oxford) and KF Thompson (Exeter) for insightful comments on the manuscript, and GW Freeman (Exeter) for species illustrations.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was funded by a Natural Environment Research Council (UK) GW4+ DTP studentship (Grant No. NE/L002434/1), Natural England (Ref. PO 904130) and the University of Exeter. Rita Castilho was funded by FCT—Foundation for Science and Technology through project UID/Multi/04326/2013. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Natural Environment Research Council (UK): NE/L002434/1.
Natural England: PO 904130.
FCT—Foundation for Science and Technology: UID/Multi/04326/2013.

Competing Interests

Rita Castilho is an Academic Editor for PeerJ.

Author Contributions

- Tom L. Jenkins conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Rita Castilho and Jamie R. Stevens conceived and designed the experiments, authored or reviewed drafts of the paper, approved the final draft.

Data Availability

The following information was supplied regarding data availability:
The raw data is available in the Supplemental File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.5684#supplemental-information.

REFERENCES

Almada F, Francisco SM, Lima CS, FitzGerald R, Mirimin L, Villegas-Ríos D, Saborido-Rey F, Afonso P, Morato T, Bexiga S, Robalo JI. 2017. Historical gene flow constraints in a northeastern Atlantic fish: phylogeography of the ballan wrasse *Labrus bergylta* across its distribution range. *Royal Society Open Science* **4**:160773 DOI 10.1098/rsos.160773.

Alvarado-Bremer JR, Vinas J, Mejuto J, Ely B, Pla C. 2005. Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. *Molecular Phylogenetics and Evolution* **36**:169–187 DOI 10.1016/j.ympev.2004.12.011.

Avise JC. 2009. Phylogeography: retrospect and prospect. *Journal of Biogeography* **36**:3–15 DOI 10.1111/j.1365-2699.2008.02032.x.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**:489–522 DOI 10.1146/annurev.es.18.110187.002421.

Barido-Sottani J, Bošková V, Plessis LD, Kühnert D, Magnus C, Mitov V, Müller NF, Pečerska J, Rasmussen DA, Zhang C, Drummond AJ, Heath TA, Pybus OG, Vaughan TG, Stadler T. 2018. Taming the BEAST—a community teaching material resource for BEAST 2. *Systematic Biology* **67**:170–174 DOI 10.1093/sysbio/syx060.

Becquet V, Simon-Bouhet B, Pante E, Hummel H, Garcia P. 2012. Glacial refugium versus range limit: conservation genetics of *Macoma balthica*, a key species in the Bay of Biscay (France). *Journal of Experimental Marine Biology and Ecology* **432–433**:73–82 DOI 10.1016/j.jembe.2012.07.008.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLOS Computational Biology* **10**:e1003537 DOI 10.1371/journal.pcbi.1003537.

Chevolot M, Hoarau G, Rijnsdorp AD, Stam WT, Olsen JL. 2006. Phylogeography and population structure of thornback rays (*Raja clavata* L., Rajidae). *Molecular Ecology* **15**:3693–3705 DOI 10.1111/j.1365-294X.2006.03043.x.

Chiverrell RC, Thomas GSP. 2010. Extent and timing of the Last Glacial Maximum (LGM) in Britain and Ireland: a review. *Journal of Quaternary Science* **25**:535–549 DOI 10.1002/jqs.1404.

Clark P, Dyke A, Shakun J, Carlson A, Clark J, Wohlfarth B, Mitrovica J, Hostetler S, McCabe M. 2009. The last glacial maximum. *Science* **325**:710–714 DOI 10.1126/science.1172873.

Correia AT, Ramos AA, Barros F, Silva G, Hamer P, Morais P, Cunha RL, Castilho R. 2012. Population structure and connectivity of the European conger eel (*Conger conger*) across the north-eastern Atlantic and western Mediterranean: integrating molecular and otolith elemental approaches. *Marine Biology* **159**:1509–1525 DOI 10.1007/s00227-012-1936-3.

Coscia I, Mariani S. 2011. Phylogeography and population structure of European sea bass in the north-east Atlantic. *Biological Journal of the Linnean Society* **104**:364–377 DOI 10.1111/j.1095-8312.2011.01712.x.

Couceiro L, Barreiro R, Ruiz JM, Sotka EE. 2007. Genetic isolation by distance among populations of the netted dog whelk *Nassarius reticulatus* (L.) along the European Atlantic coastline. *The Journal of Heredity* **98**:603–610 DOI 10.1093/jhered/esm067.

Couceiro L, López L, Sotka EE, Ruiz JM, Barreiro R. 2012. Molecular data delineate cryptic *Nassarius* species and characterize spatial genetic structure of *N. nitidus*. *Journal of the Marine Biological Association of the United Kingdom* **92**:1175–1182 DOI 10.1017/S0025315411000816.

Coyer JA, Peters AF, Stam WT, Olsen JL. 2003. Post-ice age recolonization and differentiation of *Fucus serratus* L. (Phaeophyceae; Fucaceae) populations in Northern Europe. *Molecular Ecology* **12**:1817–1829 DOI 10.1046/j.1365-294X.2003.01850.x.

Crandall ED, Sbrocco EJ, DeBoer TS, Barber PH, Carpenter KE. 2012. Expansion dating: calibrating molecular clocks in marine species from expansions onto the Sunda Shelf following the Last Glacial Maximum. *Molecular Biology and Evolution* **29**:707–719 DOI 10.1093/molbev/msr227.

Cuveliers EL, Larmuseau MHD, Hellemans B, Verherstraeten SLNA, Volckaert FAM, Maes GE. 2012. Multi-marker estimate of genetic connectivity of sole (*Solea solea*) in the North-East Atlantic Ocean. *Marine Biology* **159**:1239–1253 DOI 10.1007/s00227-012-1905-x.

D'Arcy J, Mirimin L, FitzGerald R. 2013. Phylogeographic structure of a protogynous hermaphrodite species, the ballan wrasse *Labrus bergylta*, in Ireland, Scotland, and Norway, using mitochondrial DNA. *ICES Journal of Marine Science* **70**:685–693 DOI 10.1093/icesjms/fst018.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* **22**:1185–1192 DOI 10.1093/molbev/msi103.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**:915–925.

Gómez A, Hughes RN, Wright PJ, Carvalho GR, Lunt DH. 2007. Mitochondrial DNA phylogeography and mating compatibility reveal marked genetic structuring and speciation in the NE Atlantic bryozoan *Celleporella hyalina*. *Molecular Ecology* **16**:2173–2188 DOI 10.1111/j.1365-294X.2007.03308.x.

Gysels ES, Hellemans B, Pampoulie C, Volckaert FAM. 2004. Phylogeography of the common goby, *Pomatoschistus microps*, with particular emphasis on the colonization of the Mediterranean and the North Sea. *Molecular Ecology* **13**:403–417 DOI 10.1046/j.1365-294X.2003.02087.x.

Halanych KM, Vodoti ET, Sundberg P, Dahlgren TG. 2013. Phylogeography of the horse mussel *Modiolus modiolus*. *Journal of the Marine Biological Association of the United Kingdom* **93**:1857–1869 DOI 10.1017/S0025315413000404.

Harpending HC. 1994. Signature of ancient population growth in a low resolution mitochondrial DNA mismatch distribution. *Human Biology* **66**:591–600.

Harrison F. 2011. Getting started with meta-analysis. *Methods in Ecology and Evolution* **2**:1–10 DOI 10.1111/j.2041-210X.2010.00056.x.

Hellberg ME. 2009. Gene flow and isolation among populations of marine animals. *Annual Review of Ecology, Evolution, and Systematics* **40**:291–310 DOI 10.1146/annurev.ecolsys.110308.120223.

Hewitt GM. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68**:87–112 DOI 10.1111/j.1095-8312.1999.tb01160.x.

Hewitt GM. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405**:907–913 DOI 10.1038/35016000.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B* **359**:183–195 DOI 10.1098/rstb.2003.1388.

Heyden S Von Der, Beger M, Toonen RJ, Juinio-meñez MA, Ravago-gotanco R, Fauvelot C, Bernardi G. 2014. The application of genetics to marine management. *Bulletin of Marine Sciences* **90**:1–36 DOI 10.5343/bms.2012.1079.

Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* **54**:291–301 DOI 10.1016/j.ympev.2009.09.016.

Ho SYW, Lanfear R, Bromham L, Phillips MJ, Soubrier J, Rodrigo AG, Cooper A. 2011. Time-dependent rates of molecular evolution. *Molecular Ecology* **20**:3087–3101 DOI 10.1111/j.1365-294X.2011.05178.x.

Ho SYW, Saarma U, Barnett R, Haile J, Shapiro B. 2008. The effect of inappropriate calibration: three case studies in molecular ecology. *PLOS ONE* **3**:e1615 DOI 10.1371/journal.pone.0001615.

Hoarau G, Coyer JA, Veldsink JH, Stam WT, Olsen JL. 2007. Glacial refugia and recolonization pathways in the brown seaweed *Fucus serratus*. *Molecular Ecology* **16**:3606–3616 DOI 10.1111/j.1365-294X.2007.03408.x.

Hofreiter M, Stewart J. 2009. Ecological change, range fluctuations and population dynamics during the Pleistocene. *Current Biology* **19**:R584–R594 DOI 10.1016/j.cub.2009.06.030.

Holland LP, Jenkins TL, Stevens JR. 2017. Contrasting patterns of population structure and gene flow facilitate exploration of connectivity in two widely distributed temperate octocorals. *Heredity* **119**:35–48 DOI 10.1038/hdy.2017.14.

Hughes ALC, Gyllencreutz R, Lohne ØS, Mangerud J, Svendsen JI. 2016. The last Eurasian ice sheets—a chronological database and time-slice reconstruction, DATED-1. *Boreas* **45**:1–45 DOI 10.1111/bor.12142.

Janko K, Lecointre G, Devries A, Couloux A, Cruaud C, Marshall C. 2007. Did glacial advances during the Pleistocene influence differently the demographic histories of benthic and pelagic Antarctic shelf fishes? Inferences from intraspecific mitochondrial and nuclear DNA sequence diversity. *BMC Evolutionary Biology* **7**:220 DOI 10.1186/1471-2148-7-220.

Jolly MT, Jollivet D, Gentil F, Thiébaut E, Viard F. 2005. Sharp genetic break between Atlantic and English Channel populations of the polychaete *Pectinaria koreni*, along the north coast of France. *Heredity* **94**:23–32 DOI 10.1038/sj.hdy.6800543.

Jolly MT, Viard F, Gentil F, Thiébaut E, Jollivet D. 2006. Comparative phylogeography of two coastal polychaete tubeworms in the Northeast Atlantic supports shared history and vicariant events. *Molecular Ecology* **15**:1841–1855 DOI 10.1111/j.1365-294X.2006.02910.x.

Jost L. 2008. G_{ST} and its relatives do not measure differentiation. *Molecular Ecology* **17**:4015–4026 DOI 10.1111/j.1365-294X.2008.03887.x.

Karl SA, Toonen RJ, Grant WS, Bowen BW. 2012. Common misconceptions in molecular ecology: echoes of the modern synthesis. *Molecular Ecology* **21**:4171–4189 DOI 10.1111/j.1365-294X.2012.05576.x.

Keenan K, Mcginnity P, Cross TF, Crozier WW, Prodohl PA. 2013. DiveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* **4**:782–788 DOI 10.1111/2041-210X.12067.

Kelly RP, Palumbi SR. 2010. Genetic structure among 50 species of the northeastern pacific rocky intertidal community. *PLOS ONE* **5**:e8594 DOI 10.1371/journal.pone.0008594.

Krakau M, Jacobsen S, Jensen KT, Reise K. 2012. The cockle *Cerastoderma edule* at northeast Atlantic shores: genetic signatures of glacial refugia. *Marine Biology* **159**:221–230 DOI 10.1007/s00227-011-1802-8.

Laakkonen HM, Strelkov P, Väinölä R. 2015. Molecular lineage diversity and inter-oceanic biogeographical history in *Hiatella* (Mollusca, Bivalvia). *Zoologica Scripta* **44**:383–402 DOI 10.1111/zsc.12105.

Lambeck K, Chappell J. 2001. Sea level change through the last glacial cycle. *Science* **292**:679–686 DOI 10.1126/science.1059549.

Lambeck K, Esat TM, Potter E-K. 2002. Links between climate and sea levels for the past three million years. *Nature* **419**:199–206 DOI 10.1038/nature01089.

Larmuseau MHD, Van Houdt JKJ, Guelinckx J, Hellemans B, Volckaert FAM. 2009. Distributional and demographic consequences of Pleistocene climate fluctuations for a marine demersal fish in the north-eastern Atlantic. *Journal of Biogeography* **36**:1138–1151 DOI 10.1111/j.1365-2699.2008.02072.x.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**:1451–1452 DOI 10.1093/bioinformatics/btp187.

Liggins L, Treml EA, Possingham HP, Riginos C. 2016. Seascape features, rather than dispersal traits, predict spatial genetic patterns in co-distributed reef fishes. *Journal of Biogeography* **43**:256–267 DOI 10.1111/jbi.12647.

Maggs C, Castilho R, Foltz D, Henzler C, Jolly M, Kelly J, Olsen J, Perez K, Stam W, Vainola R, Viard F, Wares J. 2008. Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. *Ecology* **89**:108–122 DOI 10.1890/08-0257.1.

Marko PB, Hoffman JM, Emme SA, McGovern TM, Keever CC, Nicole Cox L. 2010. The "Expansion-Contraction" model of pleistocene biogeography: rocky shores suffer a sea change? *Molecular Ecology* **19**:146–169 DOI 10.1111/j.1365-294X.2009.04417.x.

Ménot G, Bard E, Rostek F, Weijers JWH, Hopmans EC, Schouten S, Sinninghe Damsté JS. 2006. Early reactivation of European rivers during the last deglaciation. *Science* **313**:1623–1625 DOI 10.1126/science.1130511.

Neiva J, Assis J, Fernandes F, Pearson GA, Serrão EA. 2014. Species distribution models and mitochondrial DNA phylogeography suggest an extensive biogeographical shift in the high-intertidal seaweed *Pelvetia canaliculata*. *Journal of Biogeography* **41**:1137–1148 DOI 10.1111/jbi.12278.

Neiva J, Pearson GA, Valero M, Serrão EA. 2012. Drifting fronds and drifting alleles: range dynamics, local dispersal and habitat isolation shape the population structure of the estuarine seaweed *Fucus ceranoides*. *Journal of Biogeography* **39**:1167–1178 DOI 10.1111/j.1365-2699.2011.02670.x.

Ni G, Li Q, Kong L, Yu H. 2014. Comparative phylogeography in marginal seas of the northwestern Pacific. *Molecular Ecology* **23**:534–548 DOI 10.1111/mec.12620.

Olsen JL, Zechman FW, Hoarau G, Coyer JA, Stam WT, Valero M, Åberg P. 2010. The phylogeographic architecture of the fucoid seaweed *Ascophyllum nodosum*: an intertidal ‘‘marine tree" and survivor of more than one glacial-interglacial cycle. *Journal of Biogeography* **37**:842–856 DOI 10.1111/j.1365-2699.2009.02262.x.

Palero F, Abelló P, Macpherson E, Gristina M, Pascual M. 2008. Phylogeography of the European spiny lobster (*Palinurus elephas*): influence of current oceanographical features and historical processes. *Molecular Phylogenetics and Evolution* **48**:708–717 DOI 10.1016/j.ympev.2008.04.022.

Paradis E. 2010. Pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**:419–420 DOI 10.1093/bioinformatics/btp696.

Patarnello T, Volckaert FAMJ, Castilho R. 2007. Pillars of Hercules: is the Atlantic-Mediterranean transition a phylogeographical break? *Molecular Ecology* **16**:4426–4444 DOI 10.1111/j.1365-294X.2007.03477.x.

Pelc RA, Warner RR, Gaines SD. 2009. Geographical patterns of genetic structure in marine species with contrasting life histories. *Journal of Biogeography* **36**:1881–1890 DOI 10.1111/j.1365-2699.2009.02138.x.

Provan J, Bennett KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution* **23**:564–571 DOI 10.1016/j.tree.2008.06.010.

Provan J, Wattier RA, Maggs CA. 2005. Phylogeographic analysis of the red seaweed *Palmaria palmata* reveals a Pleistocene marine glacial refugium in the English Channel. *Molecular Ecology* **14**:793–803 DOI 10.1111/j.1365-294X.2005.02447.x.

R Core Team. 2016. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. *Available at https://www.R-project.org/*.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67(5)**:901–904 DOI 10.1093/sysbio/syy032.

Ramos-Onsins SE, Rozas J. 2002. Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19**:2092–2100 DOI 10.1093/molbev/msl052.

Remerie T, Vierstraete A, Weekers PHH, Vanfleteren JR, Vanreusel A. 2009. Phylogeography of an estuarine mysid, *Neomysis integer* (Crustacea, Mysida), along the north-east Atlantic coasts. *Journal of Biogeography* **36**:39–54 DOI 10.1111/j.1365-2699.2008.01970.x.

Riginos C, Douglas KE, Jin Y, Shanahan DF, Treml EA. 2011. Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography* **34**:566–575 DOI 10.1111/j.1600-0587.2010.06511.x.

Robalo JI, Castilho R, Francisco SM, Almada F, Knutsen H, Jorde PE, Pereira AM, Almada VC. 2012. Northern refugia and recent expansion in the North Sea: the case of the wrasse *Symphodus melops* (Linnaeus, 1758). *Ecology and Evolution* **2**:153–164 DOI 10.1002/ece3.77.

Roman J, Palumbi SR. 2004. A global invader at home: population structure of the green crab, *Carcinus maenas*, in Europe. *Molecular Ecology* **13**:2891–2898 DOI 10.1111/j.1365-294X.2004.02255.x.

Selkoe KA, Toonen RJ. 2011. Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series* **436**:291–305 DOI 10.3354/meps09238.

Siegel DA, Kinlan BP, Gaylord B, Gaines SD. 2003. Lagrangian descriptions of marine larval dispersion. *Marine Ecology Progress Series* **260**:83–96 DOI 10.3354/meps260083.

Sotelo G, Morán P, Fernández L, Posada D. 2008. Genetic variation of the spiny spider crab *Maja brachydactyla* in the northeastern Atlantic. *Marine Ecology Progress Series* **362**:211–223 DOI 10.3354/meps07433.

Taberlet P, Fumagalli L, Wust-Saucy A, Cossons J. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7**:453–464 DOI 10.1046/j.1365-294x.1998.00289.x.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**:585–595.

Toonen RJ, Andrews KR, Baums IB, Bird CE, Concepcion GT, Daly-Engel TS, Eble JA, Faucci A, Gaither MR, Iacchei M, Puritz JB, Schultz JK, Skillings DJ, Timmers MA, Bowen BW. 2011. Defining boundaries for ecosystem-based management: a multispecies case study of marine connectivity across the Hawaiian Archipelago. *Journal of Marine Biology* **2011**:1–13 DOI 10.1155/2011/460173.

Weersing K, Toonen R. 2009. Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series* **393**:1–12 DOI 10.3354/meps08287.

Weir BS, Cockerham CC. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**:1358–1370 DOI 10.2307/2408641.

Marine Policy 94 (2018) 165–173

Contents lists available at ScienceDirect

Marine Policy

journal homepage: www.elsevier.com/locate/marpol

Assessing connectivity between MPAs: Selecting taxa and translating genetic data to inform policy

Tom L. Jenkins, Jamie R. Stevens*

Department of Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK

ARTICLE INFO

Keywords:
Connectivity
Conservation genetics
Flagship species
Gene flow
Genetic connectivity
Genomics
Marine protected area networks
Marine conservation zones
Population structure

ABSTRACT

Connectivity is frequently cited as a vital component of Marine Protected Area (MPA) networks and was formally identified as one of five key principles for marine network design in European waters. Yet, without the ability to demonstrate connectivity, it is impossible to be certain that sites designated within a MPA network do in fact constitute a network, when they may –irrespective of the diversity and rarity of the taxa within them– be in reality a set of unlinked habitats and associated species assemblages. However, the process of assessing connectivity between MPAs, and which taxa to include in assessments of connectivity, is often difficult and can be dependent on a variety of factors that can be outside the control of managers, stakeholders and policymakers. Among the many methods that have been used to assess connectivity, genetic markers are often used to infer connectivity indirectly by estimating the degree of genetic differentiation between populations of a species or by inferring the origin(s) of migrants using assignment methods. While modern molecular methods can be extremely robust and are now routinely used to address conservation issues, genetic data are, to the authors' knowledge, rarely used to inform designation of MPA networks. In this paper, several biological and methodological factors are highlighted, consideration of which may help to inform the selection of species for assessments of connectivity between MPAs in a network, and this paper suggests ways in which genetic data may be interpreted to inform MPA design and policy.

1. Introduction

Connectivity is identified as a key component in the design of European Marine Protected Area (MPA) networks [1]. However, changes to the definition of connectivity outlined in many different reports [1–3] suggest there is potential confusion or conflict amongst stakeholders and scientists concerning the exact definition and function of connectivity in the context of MPA networks. The most simplistic definition is taken from Palumbi [4] whereby "*connectivity is the extent to which populations in different parts of a species range are linked by the movement of eggs, larvae or other propagules, juveniles or adults*" [1]. In contrast, other reports have outlined a more detailed definition such that maintaining connectivity involves creating "*…ecologically connected and functional networks with 'corridors or 'stepping stones' that facilitate the range shifts of populations and the movements of individuals and genes in response to ocean climate change*" [2], or that "*…the MPA network is well distributed in space and takes into account the linkages between marine ecosystems*" [3].

Connectivity is a fundamental component of population dynamics, interacting with many processes crucial to the persistence of established populations and the (re)colonisation of new habitats [5]. The study of intra-species connectivity enables the quantification of effective larval dispersal and migration between populations, while also allowing the degree of self-recruitment within populations to be estimated [6]. This is important for optimising the location and size of MPAs to create a well-connected network (instead of individual unrelated MPAs) [7,8], and for evaluating the impacts of resource exploitation on the population dynamics of commercial marine species [9].

To assess connectivity, an ideal scenario might incorporate multiple sources of data informing on connectivity from many types of taxa within the boundaries of an MPA network; however, this is often impossible due to financial and logistical constraints. Instead, managers of MPAs have typically concentrated their efforts on species that are endangered or rare, and which may be on the brink of extirpation in parts of their range, or on so-called 'umbrella', 'keystone' or 'flagship' species [10,11]. The concept of an 'umbrella' species, a species whose protection indirectly protects many other species in an ecological community, is generally recognised as appealing for assessing connectivity. This is because the establishment of a network based on such data may extrapolate the benefits of preserving the connectivity of one focal species

* Corresponding author.
E-mail address: J.R.Stevens@exeter.ac.uk (J.R. Stevens).

https://doi.org/10.1016/j.marpol.2018.04.022
Received 17 November 2017; Received in revised form 26 March 2018; Accepted 17 April 2018

Table 1
Factors to consider before selecting a species to assess genetic connectivity between Marine Protected Areas.

Type	Factor	Description	Example	Significance
Biological	Ecological importance	Does the organism have a fundamental importance to a functioning ecosystem?	- Ecosystem engineers (e.g. mussel beds, mearl)	- These species may be protected under legislation. - Higher potential impact.
	'Flagship' species	Is the species charismatic, well known to the public, and a poster child of conservation campaigns?	- Many large megafauna including cetaceans and sharks - Some threatened invertebrates	- Greater public awareness / interest. - Higher potential impact.
	Economic importance	Is the species commercially exploited?	- Fish and invertebrate coastal fisheries	- Opportunities for collaboration during sample collection.
	Taxonomy	Is the taxonomy not well resolved and the organism hard to identify?	- Sister species with very similar morphology - Very small organisms	- Species difficult to ID morphologically demand more resources and time. In some cases, a taxonomist or DNA barcoding method may be required for validation.
	Biological knowledge	A sound knowledge of the biology and ecology of the organism.	- Habitat / distribution - Larval development - Dispersal	- Improve the testing of hypotheses. - Improve interpretation of the results.
Methodological	Sample collection	Collecting tissue samples from the organism for genetic analysis.	- Feasibility / cost of collecting samples	- Protected species may require permits for tissue removal. - Logistical barriers may limit sample collection in some areas (e.g. deep sea).
	Sample sites	The number of sampling sites and the spatial separation between sites.	- Consider the number of sites needed in and around MPAs	- Adequate sampling sites in and out of MPAs could enhance hypothesis testing.
	DNA extraction	Extracting genomic DNA for analysis.	- Consider tissue type and extraction protocol before sampling	- High quantity and quality DNA can be difficult to extract from some organisms and tissue types (e.g. crustacean exoskeleton) using standard kits. - Experimentation with DNA extraction methods may be necessary, particularly for marine invertebrates.
	Sample sizes	The number of individuals per sampling site.	- Consider the number of individuals needed to draw robust conclusions	- The 'correct' number of individuals can be influenced by the choice of genetic marker. - During data analysis, the power of the markers and sample sizes can be tested using various software.
	Choice of genetic marker	Choosing a genetic marker that is polymorphic enough to investigate genetic patterns.	- Microsatellites - SNPs	- The choice and number of markers will depend on the power and resolution required.
	Availability of genetic markers	Are panels of molecular markers already available for the organism?	- Microsatellite / SNP panels	- This would avoid the need to develop markers *de novo*.

to other species in a community with similar life histories and dispersal traits. Hypothetically, a species associated with all three concepts ('umbrella', 'keystone' or 'flagship') would likely be the 'holy grail' species for studying connectivity between MPAs; however, identification of such species (if indeed they exist) has continued to elude those involved in marine conservation. Moreover, for a variety of reasons (Table 1), the study of species that come close to satisfying the criteria of a 'holy grail' species may not be feasible and, therefore, compromises are needed to facilitate the collection of data that are informative about connectivity in a given system.

1.1. Population genetic structure

Genetic markers are commonly used in ecology to study the spatial genetic structure of a species. Such data can facilitate exploration of patterns of genetic diversity, and can enable researchers to detect genetic differences between samples and to ascertain whether the organisms at each sampling site constitute a discrete population. However, identifying discrete populations can be challenging in marine species [12] and, while low genetic differentiation between populations may imply high gene flow (or large effective population sizes), the same pattern could also suggest that multiple samples are from a single panmictic population. Thus, when using population genetic structure to infer patterns of connectivity, it is important to collect spatially discrete samples and to be aware of the challenges of defining populations.

Three main evolutionary processes influence the population genetic structure of a species across space and time: gene flow, genetic drift and natural selection [13,14]. Low genetic differentiation between populations may be driven by high gene flow because the transfer of genetic material homogenises allele frequencies. However, large effective population sizes may also result in low genetic differentiation between populations, though, again, accurate estimation of effective population size in marine organisms can be extremely difficult [15]. In the absence of gene flow, allele frequencies can diverge over time because of the random sampling of alleles from generation to generation (genetic drift). Genetic drift is stronger in small or bottlenecked populations because sampling variance is greater when effective population size is smaller [16]. Populations can also diverge when strong natural selection favours a particular mutation that increases the fitness or survivorship of the carriers, resulting in the allele sweeping to fixation in that particular population. For studies where the primary goals are to assess inbreeding, effective population size(s) or connectivity, researchers have commonly employed neutral markers because genetic patterns at these markers are driven by the interacting processes of gene flow and genetic drift, and not selection [17]. However, hitchhiking neutral markers and markers under selection have been shown to sometimes provide more power for directly tracking migrants in assignment studies [18]; as the use of non-neutral markers in molecular ecology increases, this will likely have promising applications for inferring patterns of connectivity [19].

1.2. Genetic connectivity

Measuring dispersal and connectivity using conventional tracking methods (e.g. physical tags, satellite telemetry) is extremely difficult in many coastal marine species because of their typically large ranges and pelagic larval phases [5]. Genetic markers are naturally present in every individual in a population and this makes them ideal to infer patterns of connectivity in such species. Genetic connectivity is defined by Lowe and Allendorf [20] as the "*degree to which gene flow affects evolutionary processes within subpopulations*". In other words, individuals must disperse to a new population and must successfully contribute their genes to the next generation to facilitate genetic connectivity. In contrast, demographic connectivity refers to how the absolute number of exchanges (via immigration or emigration) between populations affects growth and vital rates within populations [20]. Studies of population genetic structure can be used to infer genetic connectivity, however, they generally provide little information about demographic connectivity, unless combined with other data such as direct estimates of dispersal or abundance [20] or biophysical modelling [21].

To assess contemporary genetic connectivity, two types of methods are generally employed: (i) indirectly inferring genetic connectivity by examining genetic similarities or dissimilarities (genetic structure) between spatially discrete populations, or (ii) directly estimating genetic connectivity by detecting migrants through population or parentage assignment [19,22,23]. For both methods, the most widely used markers are microsatellites and single nucleotide polymorphisms (SNPs). However, SNPs are fast becoming the marker system of choice, particularly in non-model organisms [24,25], because the rapid advancement of high-throughput sequencing methodologies [26,27] can enable thousands to tens of thousands of markers to be discovered and genotyped, as opposed to tens of markers using microsatellites. Moreover, difficulties in cross-calibrating microsatellite allele sizes between sequencing platforms and laboratories [28] has limited their use in broad-scale studies, a limitation which does not affect SNPs. The use of genomic SNPs therefore provides wider coverage across the genome and potentially greater power for resolving patterns of population structure and genetic connectivity at finer spatial scales [29].

However, the general lack of genetic evidence used by marine policymakers and managers suggests much of the genetic/genomic data generated are currently not considered during the planning and designation of MPAs. In this paper, a number of biological and methodological factors are highlighted that should be considered before selecting taxa to assess genetic connectivity between MPAs. In addition, using published data from a previous study (Holland *et al.* [32]) and the MPA network in southwest Britain as an example, this paper discusses how genetic data from a typical population genetic/genomic study may be interpreted to inform managers about connectivity in a MPA network, and which areas to consider prioritising to maximise the protection of biodiversity.

2. Selecting taxa

The selection of appropriate taxa to use as surrogates for assessments of genetic connectivity between MPAs has seldom been discussed in the literature (but see Marti-Puig et al. [30]). Coastal benthic marine invertebrates are often good candidates because they can be relatively abundant with large ranges, and dispersal is typically defined during a pelagic phase undertaken by an early life stage (e.g. eggs or larvae), while the adults remain relatively sedentary [5]. This type of development means connectivity is mainly dependent on local hydrological conditions (as well as species-specific traits) and, therefore, better reflects natural patterns of connectivity, as opposed to studying connectivity driven by organismal behaviour in motile and migratory species. Since patterns of genetic connectivity can vary between species over similar geographical areas [31,32], it is important to consider assessing connectivity in more than one species with differing biology/ecology. This allows the exploration of species-specific genetic connectivity and patterns of connectivity common across taxa to be identified [30].

2.1. Biological factors

Some biological features of candidate species can inevitably enhance the public appeal and societal impact of a study, while other features can limit the collection of samples and the interpretation of data generated by genetic markers (Table 1). For the purpose of promoting marine conservation, charismatic megafauna such as marine mammals and sharks frequently dominate awareness campaigns ('flagship' species) because they can raise funds and change public opinions and behaviour. Although many of these species may not be the best candidates for assessing MPA connectivity, these enigmatic

animals are typically well-known by the wider public and benefit from a greater awareness and potential impact than other marine fauna. As a result, if a candidate species is poorly known to the public community, highlighting its importance for the conservation of an associated enigmatic species may have an equivalent effect (e.g. the interactions between kelp forests and sea otters [33]).

Benthic marine invertebrates are generally not 'flagship' species (though there are exceptions, e.g. pink sea fans). However, it is recognised that many benthic invertebrates have a crucial ecological role (e.g. mussel beds as ecosystem engineers / habitat builders) or are commercially exploited (e.g. European lobsters), meaning they are either fundamentally important to the ecosystem or the local/regional economy, or both. This may encourage relevant management bodies and/or stakeholders to collaborate, to contribute funding and/or to share equipment (depending on the organisation's interests and capacity), all of which can serve to advance a particular project. For example, lobster fishermen have access to a potential myriad of individuals from which tissue samples can be obtained. Forming these types of collaborations can facilitate access to a virtually unlimited number of samples depending on the fishery status, thereby avoiding the need to arrange dedicated sampling trips, and the associated costs and researcher time typically required for collection. Moreover, maintaining dialogue with such a stakeholder(s) may promote more effective communication of the potential benefits of the research and, ultimately, dissemination of the results.

Other factors to consider include whether the biology and ecology of the candidate species are well known. This process starts –perhaps obviously– by accurate identification of the candidate species and avoiding the erroneous inclusion of closely related or cryptic species, which can drastically influence the results of population genetic structure analyses [34]. The difficulty of accurate taxonomic identification can be further exacerbated when the organism is very small; in some cases, a second opinion from a dedicated taxonomist or molecular verification (e.g. DNA barcoding) may be required. In addition, a thorough understanding of the dispersal, life history and habitat of the candidate species will usually help to explain some of the genetic patterns observed, thereby improving interpretation of the genetic data.

2.2. Methodological factors

The sampling design of a study should be carefully considered prior to sample collection to ensure that the resulting genetic data are robust and applicable for assessments of genetic connectivity. This typically includes assessing whether the desired sampling strategy is feasible and that sufficient tissue samples from a broad enough range of sites can be taken for meaningful genetic analysis. For example, as suggested previously, if an organism is commercially fished, it may be possible to have tissue samples collected *in situ* by fishery personnel. Moreover, ensuring that samples of a species of interest are collected from both within the boundaries of a MPA network and from sites outside ensures that hypotheses about connectivity beyond MPA boundaries can be tested. This approach has provided useful data in several previous studies [32,35,36], allowing the performance of a MPA network to be evaluated for the species being studied.

Other factors to consider include the type(s) of tissue to sample and which genetic markers to use in assessments of population genetic structure. This is of critical importance because the type of tissue can profoundly influence the quantity and quality of DNA obtained post-extraction. For example, crustacean exoskeletal tissues, such as pleopods, are advantageous because they are easily obtained and constitute a non-destructive tissue sample; however, extracting sufficient amounts of pure (contaminant-free) DNA from these tissue types can be extremely difficult using both conventional and kit-based protocols [37]. Moreover, obtaining high molecular weight, non-degraded DNA can be important for methods that utilise next-generation sequencing technology, for example, whole-genome sequencing and SNP discovery from restriction-site associated DNA sequencing (RADseq) [38–40]. In these cases, optimising the preservation and extraction of DNA will need to be considered prior to sampling and DNA extraction. Choosing appropriate genetic markers and the method of isolation for studies of population genetic structure is also a non-trivial task. Discussion of which genetic markers to employ for a particular study is outside the scope of this paper; however, a number of comprehensive review papers have been published to address this question [13,41–44]. In addition, tools exist that can help practitioners choose the appropriate number of samples and genetic markers (e.g. SPOTG [45]). Prior to commencing development work, the literature should be screened thoroughly to determine whether genetic markers of a suitable resolution are already available for a candidate species – this can avoid the costs and time typically required for the development of novel markers. For example, SNP panels are now available for a wide range of marine species (e.g. salmonids [46], crustaceans [47] and molluscs [48]), and are likely to be useful for the analysis of genetic structure, population assignment and connectivity.

3. Translating genetic data to inform policy

Translating primary research into the language and terminology required by policymakers and conservation managers to allow them to make decisions is not a trivial task. Often, it may be more beneficial to present a few points that represent the key findings of a study, while trying to avoid unnecessary technical jargon, which could lead to misinterpretation or confusion. Several papers have discussed the challenges of translating genetic data to inform management and have asserted the importance of strong collaboration and communication between scientists and practitioners [49–55]. Some of the reasons put forward for the avoidance of genetic data in fisheries management include a lack of understanding of the potential value of genetic data, the assumption that genetic studies are expensive, and the suggestion that other data types are significantly more important than genetic information in management decisions [9]. One feature of genetic data is that they cannot be seen or measured without the use of specialist molecular techniques, meaning it can sometimes be difficult to articulate the level of variation and the importance of genetic diversity to non-scientists [49]. Moreover, in cases where research is carried out by non-academic bodies, these institutions often have little incentive to publish, or have internal deadlines or political/legal constraints that may delay scientific publication, so the findings may not be widely-disseminated [52].

However, while some barriers to the dissemination of genetic research exist, there are examples across various taxa and systems where genetic data have successfully informed policy and conservation, and have led to improved management decisions. This suggests that some barriers to the application of genetic data are starting to be overcome. Some examples include the genetic restoration of Florida panthers [56], the selective reintroduction of endangered Burmese roof turtles [57], the genetic management of salmonids [58,59], the identification of stock/management units for commercial species in the Mediterranean [60], the authenticity and monitoring of seafood in sushi bars [61], and the traceability of fisheries resources (e.g. FishPopTrace, [62]). In the latter case, the FishPopTrace Consortium was an international project funded by the European Union with the aim of developing genetic marker panels capable of pinpointing the stock/population origin of a particular individual from a species [18,62]. Monitoring the origin of a fishery product is seen as a strategy to potentially increase transparency in the food supply chain and reduce illegal, unreported and unregulated fishing (IUU) and product mislabelling [62]. The project focused on four commercially important fish species: cod (*Gadus morhua*), hake (*Merluccius merluccius*), herring (*Clupea harengus*) and common sole (*Solea solea*), and the results indicated that gene-associated SNP markers could assign individual fish to their population of origin at 93–100% accuracy across a range of spatial scales [18]. The study

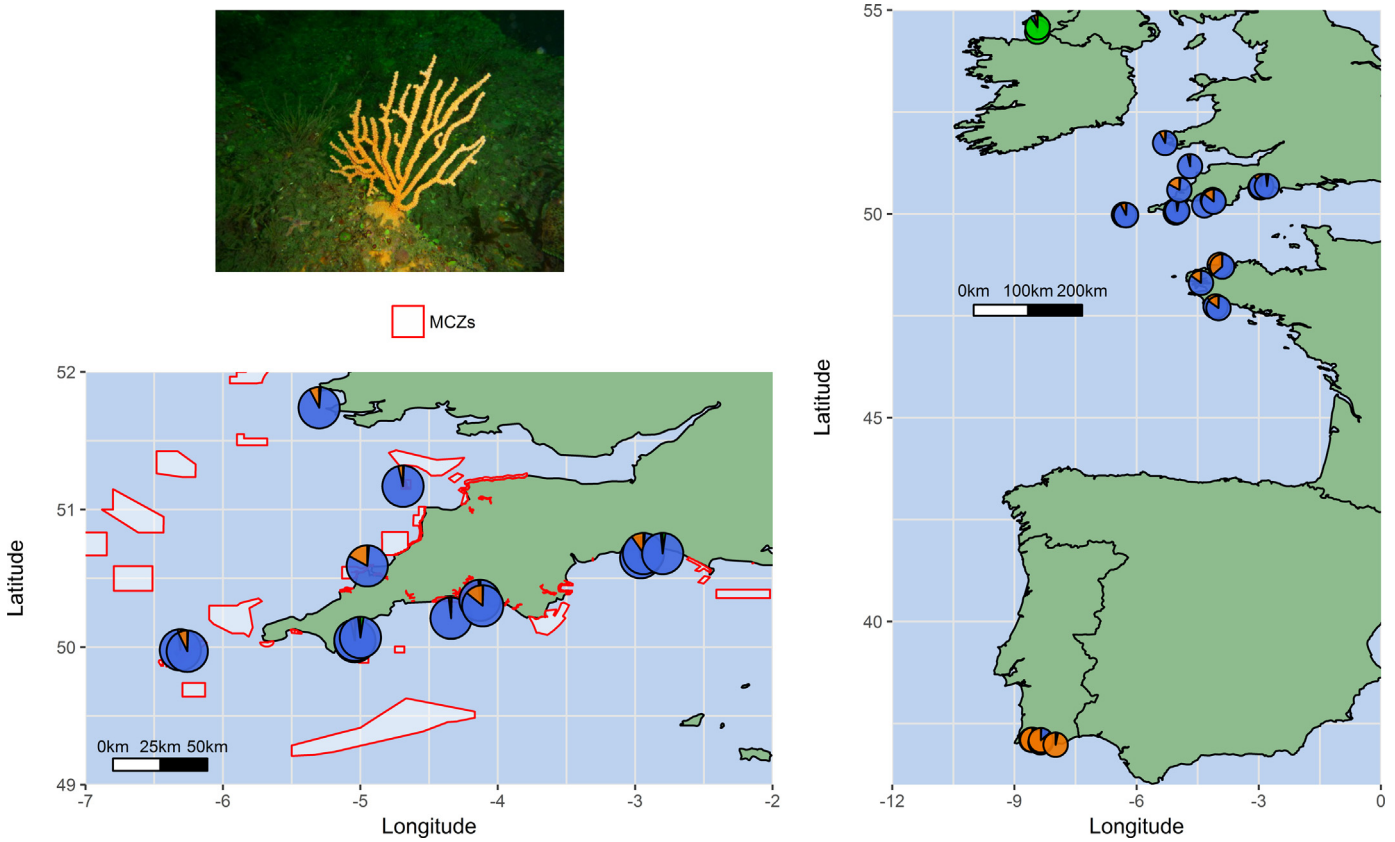


Fig. 1. Pink sea fan (*Eunicella verrucosa*) (top left) average cluster memberships derived from STRUCTURE analysis of 13 microsatellite markers – re-drawn from Holland et al. [32]. The right map shows the average membership coefficients for each genetic cluster for each population studied in the original paper. The bottom left map zooms in on southwest Britain and includes overlays (red outlines) of the Marine Conservation Zones designated in English and Welsh waters to-date.

illustrates the remarkable potential of genetic data to help enforce fisheries regulations and conservation measures across different species and geographical areas.

Yet, while there are a myriad of studies documenting the spatial genetic structure and genetic connectivity of benthic marine species, very few of these studies, to our knowledge, have been directly used as evidence to inform or support MPA designations and/or network connectivity. This may be a consequence of ineffective dissemination of the key findings of research projects to managers and policymakers, but also likely relates to the availability of data at the time when large-scale MPA projects were commissioned and candidate lists were first drawn-up. Nevertheless, as these data are becoming more available to practitioners, it is crucial that gaps between primary research (i.e. academic researchers) and applied science (i.e. policymakers) are overcome in order to realise the potential of genetic data to inform MPA design and conservation planning [50,52].

3.1. UK MPA network

Across the UK, as of December 2017, approximately 23% of marine/estuarine environments are within MPAs (see: http://jncc.defra.gov.uk/page-4549). At the time of writing, the network comprised 298 MPAs including: Special Areas of Conservation (SACs; 105) and Special Protected Areas (SPAs; 106) with marine components, Nature Conservation Marine Protected Areas (NCMPAs; 30 in Scotland), Ramsar sites (Isle of Man), and Marine Conservation Zones (MCZs; 56 in England, Wales and Northern Ireland). This network of MPAs has been created to satisfy the UK's commitments to the Convention on Biological Diversity, EU habitat and marine strategy regulations, and the Oslo/Paris (OSPAR) Convention to protect the marine environment of the northeast Atlantic.

In England and Wales, the MCZ project began in 2008 with the aim of filling gaps in the MPA network and potentially addressing any deficits in connectivity; following the EU Marine Strategy Framework Directive (2008), the UK Marine and Coastal Access Act (2009) enshrined the designation of MCZs into UK Law. After identification of 127 candidate MCZs in 2011 by four regional stakeholder groups, 50 MCZs in England (27 in tranche one, November 2013; 23 in tranche two, January 2016) and one in Wales (Skomer Island, Pembrokeshire, 2014) have been designated, with a final tranche for England to be announced in 2018. This has, in the view of some commentators, coincided with a shift from a bottom-up to a top-down approach, with stakeholder engagement now limited to bilateral consultations [63]. The MCZ project has also steered away from its initial focus on broad-scale networks and instead has concentrated efforts on single-feature conservation [63], such as protecting vulnerable species (e.g. pink sea fans) and key habitats (e.g. intertidal boulder communities).

One of the main ambitions of the UK MPA project was to create an ecologically coherent network of MPAs, for which connectivity was seen as one of five key planning principles, alongside representativity, adequacy, replication, and ecologically and biologically significant areas [64]. Assessing connectivity of the English and Welsh MPA network has primarily focused on linking discrete habitats (e.g. littoral rock and hard substrata, sublittoral sediment, etc.), such that each habitat is represented by a MPA every 80 km or less [3,65], the spacing recommended by Roberts et al. [66] to maintain ecological connectivity. Connectivity for a discrete habitat is deemed sufficient when 40 km buffers drawn around two adjacent MPAs converge [3,65]. For many benthic marine species, defining a network in this way may be sufficient to maintain connectivity between nearby populations. However, it is important to note that it may not suit all species because connectivity can be influenced by a number of biological (e.g. larval

Table 2
Summary of the promises and pitfalls of genetic data for informing Marine Protected Area design.

Genetic data	Description	Promises	Pitfalls
Genetic diversity	The amount of genetic variation contained within a population or species. Statistic is a combination of the number of allelic variants and their frequency in a sample.	• Populations with unique/high genetic diversity may have more resilience to environmental change. • Can inform the location and boundaries of MPAs. • Could prioritise placement of MPAs to safeguard this diversity.	• Patterns and magnitude of genetic diversity measures may differ depending on the molecular marker used. • Mutations in primer sites can lead to null alleles which can lead to inaccurate estimates of genetic diversity.
Population genetic structure	The spatial distribution of genetic variation among populations in a species, allowing genetic similarities or dissimilarities between sample groups to be explored.	• Infer gene flow between populations. • Identify potentially genetically isolated and source/sink populations. • Indirectly infer dispersal distances and genetic connectivity. • Can inform the location and boundaries of MPAs. • Infer connectivity between MPAs, providing information about ecological coherency.	• Large effective population size, not gene flow, can be responsible for low genetic differentiation. • It is difficult to infer gene flow from marine species with overlapping or long-life spans. • Genetic markers only provide information on the number of effective migrants. Genetic data cannot reliably estimate demographic connectivity without additional data. • A single effective migrant per generation can homogenise populations; thus, genetically similar populations may have only very limited larval exchange.
Population assignment (Individual assignment)	Assign an individual to a population or cluster in which their genotype has the highest probability of occurring.	• Infer the origin of an individual and track migrants. • Infer dispersal distances and genetic connectivity. • Infer connectivity between MPAs, providing information about ecological coherency.	• Requires sound knowledge of the species distribution. • Accuracy reduces with decreasing genetic differentiation. Therefore, markers with high power to distinguish differences are necessary for species with low overall levels of genetic differentiation.
Parentage assignment	Assign an individual to their biological parents based on their genotypes.	• Infer the origin of an individual and track migrants. • Infer dispersal distances and genetic connectivity. • Infer connectivity between MPAs, providing information about ecological coherency.	• Requires sound knowledge of the species distribution. • Requires a significant proportion of potential parents to be sampled. Can be logistically difficult to sample a sufficient proportion of contributing parents to make assignment accurate.

dispersal [36], spawning periodicity [67]) and hydrological (e.g. ocean currents and fronts [68]) factors, which further complicate the positioning of MPAs within a network.

3.2. Genetic data as evidence: pink sea fan case study

Genetic data are currently not (to the authors' knowledge) used by managers as evidence to inform MPA designation or network connectivity in England and Wales. Discussions with national agencies suggest that the personnel and infrastructure are not in place to process, grade and assess the usefulness of spatially relevant genetic data. This may explain the lack of genetic data currently used as evidence to support existing MPA designations or to inform new designations around southwest Britain. However, genetic data from single-species studies can provide an empirical estimate of connectivity within evolutionary timescales [20]. This, therefore, gives an approximation of genetic connectivity over the last few generations in the species studied [50], which would likely supplement the present assessments of connectivity discussed in Section 3.1. Moreover, genetic data can reveal distinct localised genetic diversity –otherwise undetectable using only presence/absence data or biophysical modelling– which can be of major importance for identifying populations or areas that should be prioritised for protection.

In Fig. 1, STRUCTURE [69] results taken from a recent study [32] are presented; the study analysed patterns of variation at 13 microsatellite loci and explored the population structure and genetic connectivity of a 'flagship' species in English and Wales, *Eunicella verrucosa* (the pink sea fan). *Eunicella verrucosa* is listed as 'Vulnerable' by the IUCN Red List and is a Biodiversity Action Plan priority species in English and Welsh waters; accordingly, several MCZs specifically identify *E. verrucosa* as a protected feature in their designations (e.g. The Manacles, The Isles of Scilly, Chesil Beach and Stennis Ledges). Moreover, colonies are sessile, dispersal is achieved by broadcast spawning, and 60% of colonies recorded by diver surveys fall within MPAs [70]; therefore, *E. verrucosa* fulfils several criteria associated with the ideal surrogate species to assess connectivity between MPAs. In Fig. 1, each pie chart represents a sampling site and the colours represent genetic cluster memberships for each population, averaged across all individuals in that population. In effect, when two pie charts are primarily composed of the same colour, this implies that these two populations are genetically similar, suggesting high genetic connectivity (or large effective population sizes). In southwest Britain, the composition of the pie charts are relatively similar, indicating genetic similarity; as the original authors report, this suggests that the current MPA network is likely sufficient to maintain genetic connectivity in this species across southwest England and Wales. In comparison, Portuguese and Irish colonies are genetically different. As reported in the original study, the genetic differences observed in Portugal likely represent a stepping-stone model of genetic connectivity, driven by isolation-by-distance, whereby gene flow occurs more frequently between populations that are closer together than further apart and over time populations diverge due to genetic drift. In contrast, the authors suggested the genetic differences observed in northwest Ireland could be the result of a barrier to gene flow and subsequent genetic drift, or possibly a result of local adaptation driven by natural selection at this northerly location, though it was unclear which process was primarily responsible [32].

The pink sea fan study discussed above was indirectly commissioned and funded by the UK Government with the aim of assessing connectivity of *E. verrucosa* using genetic techniques, and with the potential to inform and support the designations of MPAs that included *E. verrucosa*. The key finding of this study which might constitute evidence

for MPA managers is that, as it stands, the MPA network in southwest Britain appears to be sufficient to maintain genetic connectivity in this protected species [32]. The integration of these data in future reviews or monitoring reports would likely serve as another piece of evidence to support the designation of these MPAs and to help demonstrate the ecological coherency of the network in southwest Britain. To facilitate more efficient translation and transparency going forward, researchers aiming to inform MPA designation using population genetic data are encouraged to create a visual representation similar to Fig. 1 to better simplify and standardise interpretation for managers and policymakers.

3.3. Promises and pitfalls of genetic data

Genetic data have much promise in informing the planning stages of MPA network design and in supporting previously designated MPAs. Accordingly, it is important that managers and policymakers are aware of the opportunities provided by genetic studies but are also aware of some of the pitfalls that are linked to the methods and interpretations before action is undertaken (Table 2).

Genetic data have the potential to inform managers in two main ways: single feature designations and network connectivity. For priority species that cover a relatively large spatial area within and across political boundaries (e.g. pink sea fans), designating a MPA has to be strategic and knowing where to place a MPA can be extremely difficult. By studying the population genetic structure, data of suitable resolution may allow managers to identify key populations or areas that harbour unique or high genetic diversity, as shown in the case study above, thereby providing additional ecological evidence to support a designation. This information may also be useful for determining the boundaries of MPAs that are designated to safeguard protected species. The appropriate authorities should be encouraged to consider protecting these rare genetic variants, particularly in ecologically or economically important species, even if the exact cause of the unique diversity is not known. This is because if these individuals/populations are wiped out due to anthropogenic causes, this diversity will be permanently lost to the species before these genetic variants have an opportunity (potentially) to benefit the species in a constantly dynamic environment. Such studies can also allow a species' effective population size to be inferred which can reveal the genetic health of a population; however, estimating effective population size in marine species can be notoriously difficult [15]. For MPA networks, information about dispersal distances, potentially isolated populations, and connectivity between populations or habitat patches can be inferred through the analysis of intraspecific genetic data [5]. Additionally, by using appropriate genetic software, the direction of gene flow can be estimated – this can be particularly important for identifying source populations that export potential recruits to nearby populations [71]. The protection of source populations is extremely important for marine conservation because it can facilitate replenishment or recolonisation of populations that have suffered from local declines and/or are not self-sustaining (i.e. sites relying on immigrants to maintain healthy population sizes) [5]. However, while genetic data can provide some understanding into source and sink dynamics on evolutionary timescales, these analyses would likely benefit from incorporating biophysical modelling data into their conclusions to evaluate whether contemporary hydrological conditions (e.g. ocean currents) are the potential driver behind any asymmetrical connectivity found. Such integrative studies have typically been referred to as 'seascape genetics/genomics' [72,73] and have been shown to enhance studies of marine connectivity by providing insights into both demographic and genetic connectivity [21,74].

Limitations associated with the inference of genetic data are usually reported in the original published studies; nonetheless, some general limitations are discussed here. Firstly, managers are typically interested in the absolute number of migrants (demographic connectivity), but genetic markers can only provide information on the number of effective migrants (e.g. individuals/larvae that successfully disperse to a new population and reproduce/survive to the next generation). Moreover, a single effective migrant per generation can be sufficient to homogenise populations [20], meaning, despite being genetically similar, some populations may have minimal larval exchange [23]. Parentage assignment can circumvent this issue to some degree by attempting to track migrants; however, to be useful, these methods require a significant proportion of potential parents to be sampled/characterised [19]. Secondly, inferring patterns of connectivity from marine species with overlapping generations or long-life spans (e.g. corals) can be difficult because genetic profiles can remain essentially unchanged for many decades, even after barriers to gene flow have been introduced [23]. Therefore, interpretations of population differentiation and genetic structure can, in some cases, represent historical and not contemporary gene flow [75]. This difference in timescales is critical to consider in assessments of connectivity because MPA networks are generally established to protect and maintain present-day and future patterns of diversity and connectivity, or to facilitate recovery/restoration to a previous level of abundance and diversity. However, as with genetic data, most methods of assessing connectivity have their own assumptions and limitations, so consideration of all of the best available scientific knowledge will be crucial to create well-connected networks that maximise the protection of marine biodiversity.

4. Conclusion

Over 11,000 MPAs have been designated globally (~3.7% of global oceans) to protect the world's oceans (http://www.mpatlas.org/explore/). There are also numerous published studies of population genetic structure for a variety of marine organisms across small (i.e. within seas) and large (i.e. across seas and oceans) geographic scales. Therefore, the potential for genetic data to provide evidence to support the designation of existing or new MPAs is profound. In this paper, a number of factors are presented that could help practitioners select appropriate taxa to assess connectivity between MPAs. In addition, this paper has discussed how genetic data from a typical population genetics/genomics study may be interpreted to inform MPA designation and network connectivity. These two sections are anticipated to be useful for managers involved in MPA designation processes, and particularly for those tasked with the designation, monitoring, review and enforcement of the current UK MPA network.

Acknowledgements

We thank Emma Dade and James Highfield (Natural England) for valuable discussions concerning this research and for critical reading of the manuscript. This research was funded by a Natural Environment Research Council UK GW4+ DTP studentship (Grant No. NE/L002434/1), Natural England (Ref. PO 904130) and the University of Exeter, and forms part of the PhD of Tom L. Jenkins.

Data accessibility

R code and data files required to produce Fig. 1 are available as Supplementary material.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.marpol.2018.04.022.

References

[1] OSPAR Commission, An assessment of the ecological coherence of the OSPAR Network of Marine Protected Areas in 2012. ISBN: 978-1-909159-52-5, 2013.

[2] OSPAR Commission, Results of the e-consultation on the draft quality status report 2010. ISBN 978-1-907390-41-8, 2010.
[3] H. Carr, A. Cornthwaite, H. Wright, J. Davies, Assessing progress towards an ecologically coherent network of MPAs in Secretary of State Waters in 2014, JNCC Report. 〈http://jncc.defra.gov.uk/PDF/JNCC_NetworkProgressInSoSWaters_2014.pdf〉.
[4] S.R. Palumbi, Population genetics, demographic connectivity, and the design of marine reserves, Ecol. Appl. 13 (2003) 146–158, http://dx.doi.org/10.1890/1051-0761(2003)013[0146:PGDCAT]2.0.CO;2.
[5] R.K. Cowen, S. Sponaugle, Larval dispersal and marine population connectivity, Ann. Rev. Mar. Sci. 1 (2009) 443–466, http://dx.doi.org/10.1146/annurev.marine.010908.163757.
[6] R.K. Cowen, K.M. Lwiza, S. Sponaugle, C.B. Paris, D.B. Olson, Connectivity of marine populations: open or closed? Science 287 (2000) 857–859, http://dx.doi.org/10.1126/science.287.5454.857.
[7] G.P. Jones, M. Srinivasan, G.R. Almany, Population connectivity and conservation of marine biodiversity, Oceanography 20 (2007) 100–111, http://dx.doi.org/10.5670/oceanog.2007.33.
[8] G.R. Almany, S.R. Connolly, D.D. Heath, J.D. Hogan, G.P. Jones, L.J. McCook, M. Mills, R.L. Pressey, D.H. Williamson, Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs, Coral Reefs 28 (2009) 339–351, http://dx.doi.org/10.1007/s00338-009-0484-x.
[9] L. Bernatchez, M. Wellenreuther, C. Araneda, D.T. Ashton, J.M.I. Barth, T.D. Beacham, G.E. Maes, J.T. Martinsohn, K.M. Miller, K.A. Naish, J.R. Ovenden, C.R. Primmer, H. Young Suk, N.O. Therkildsen, R.E. Withler, Harnessing the power of genomics to secure the future of seafood, Trends Ecol. Evol. 32 (2017) 665–680, http://dx.doi.org/10.1016/j.tree.2017.06.010.
[10] D. Simberloff, Flagships, umbrellas, and keystones: is single-species management passé in the landscape era? Biol. Conserv. 83 (1998) 247–257 〈http://www.sciencedirect.com/science/article/pii/S0006320797000815)〉.
[11] G. Kalinkat, J.S. Cabral, W. Darwall, G.F. Ficetola, J.L. Fisher, D.P. Giling, M.P. Gosselin, H.P. Grossart, S.C. Jähnig, J.M. Jeschke, K. Knopf, S. Larsen, G. Onandia, M. Pätzig, W.C. Saul, G. Singer, E. Sperfeld, I. Jarić, Flagship umbrella species needed for the conservation of overlooked aquatic biodiversity, Conserv. Biol. 31 (2017) 481–485, http://dx.doi.org/10.1111/cobi.12813.
[12] R.S. Waples, O. Gaggiotti, What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity, Mol. Ecol. 15 (2006) 1419–1439, http://dx.doi.org/10.1111/j.1365-294X.2006.02890.x.
[13] M. Hellberg, R. Burton, Genetic assessment of connectivity among marine populations, Bull. Mar. Sci. 70 (2002) 273–290 〈http://www.ingentaconnect.com/content/umrsmas/bullmar/2002/00000070/A00101s1/art00003〉.
[14] M.E. Hellberg, Gene flow and isolation among populations of marine animals, Annu. Rev. Ecol. Evol. Syst. 40 (2009) 291–310, http://dx.doi.org/10.1146/annurev.ecolsys.110308.120223.
[15] M.P. Hare, L. Nunney, M.K. Schwartz, D.E. Ruzzante, M. Burford, R.S. Waples, K. Ruegg, F. Palstra, Understanding and estimating effective population size for practical application in marine species management, Conserv. Biol. 25 (2011) 438–449, http://dx.doi.org/10.1111/j.1523-1739.2010.01637.x.
[16] B. Charlesworth, Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation, Nat. Rev. Genet. 10 (2009) 195–205, http://dx.doi.org/10.1038/nrg2526.
[17] W.C. Funk, J.K. McKay, P.A. Hohenlohe, F.W. Allendorf, Harnessing genomics for delineating conservation units, Trends Ecol. Evol. 27 (2012) 489–496, http://dx.doi.org/10.1016/j.tree.2012.05.012.
[18] E.E. Nielsen, A. Cariani, E. Mac Aoidh, G.E. Maes, I. Milano, R. Ogden, M. Taylor, J. Hemmer-Hansen, M. Babbucci, L. Bargelloni, D. Bekkevold, E. Diopere, L. Grenfell, S. Helyar, M.T. Limborg, J.T. Martinsohn, R. McEwing, F. Panitz, T. Patarnello, F. Tinti, J.K.J. Van Houdt, F.A.M. Volckaert, R.S. Waples, J.E.J. Albin, J.M. Vieites Baptista, V. Barmintsev, J.M. Bautista, C. Bendixen, J.-P. Bergé, D. Blohm, B. Cardazzo, A. Diez, M. Espiñeira, A.J. Geffen, E. Gonzalez, N. González-Lavín, I. Guarniero, M. Jeráme, M. Kochzius, G. Krey, O. Mouchel, E. Negrisolo, C. Piccinetti, A. Puyet, S. Rastorguev, J.P. Smith, M. Trentini, V. Verrez-Bagnis, A. Volkov, A. Zanzi, G.R. Carvalho, Gene-associated markers provide tools for tackling illegal fishing and false eco-certification, Nat. Commun. 3 (2012) 851, http://dx.doi.org/10.1038/ncomms1845.
[19] P.-A. Gagnaire, T. Broquet, D. Aurelle, F. Viard, A. Souissi, F. Bonhomme, S. Arnaud-Haond, N. Bierne, Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era, Evol. Appl. 8 (2015) 769–786, http://dx.doi.org/10.1111/eva.12288.
[20] W.H. Lowe, F.W. Allendorf, What can genetics tell us about population connectivity? Mol. Ecol. 19 (2010) 3038–3051, http://dx.doi.org/10.1111/j.1365-294X.2010.04688.x.
[21] L. Thomas, J.J. Bell, Testing the consistency of connectivity patterns for a widely dispersing marine species, Heredity 111 (2013) 345–354, http://dx.doi.org/10.1038/hdy.2013.58.
[22] A.G. Jones, C.M. Small, K.A. Paczolt, N.L. Ratterman, A practical guide to methods of parentage analysis, Mol. Ecol. Resour. 10 (2010) 6–30, http://dx.doi.org/10.1111/j.1755-0998.2009.02778.x.
[23] L.W. Botsford, J.W. White, M.A. Coffroth, C.B. Paris, S. Planes, T.L. Shearer, S.R. Thorrold, G.P. Jones, Connectivity and resilience of coral reef metapopulations in marine protected areas: matching empirical efforts to predictive needs, Coral Reefs 28 (2009) 327–337, http://dx.doi.org/10.1007/s00338-009-0466-z.
[24] S.J. Helyar, J. Hemmer-Hansen, D. Bekkevold, M.I. Taylor, R. Ogden, M.T. Limborg, A. Cariani, G.E. Maes, E. Diopere, G.R. Carvalho, E.E. Nielsen, Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges, Mol. Ecol. Resour. 11 (2011) 123–136, http://dx.doi.org/10.1111/j.1755-0998.2010.02943.x.
[25] J.E. Seeb, G. Carvalho, L. Hauser, K. Naish, S. Roberts, L.W. Seeb, Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms, Mol. Ecol. Resour. 11 (2011) 1–8, http://dx.doi.org/10.1111/j.1755-0998.2010.02979.x.
[26] K.R. Andrews, J.M. Good, M.R. Miller, G. Luikart, P.A. Hohenlohe, Harnessing the power of RADseq for ecological and evolutionary genomics, Nat. Rev. Genet. 17 (2016) 81–92, http://dx.doi.org/10.1038/nrg.2015.28.
[27] J.W. Davey, P.A. Hohenlohe, P.D. Etter, J.Q. Boone, J.M. Catchen, M.L. Blaxter, Genome-wide genetic marker discovery and genotyping using next-generation sequencing, Nat. Rev. Genet. 12 (2011) 499–510, http://dx.doi.org/10.1038/nrg3012.
[28] J.S. Ellis, J. Gilbey, A. Armstrong, T. Balstad, E. Cauwelier, C. Cherbonnel, S. Consuegra, J. Coughlan, T.F. Cross, W. Crozier, E. Dillane, D. Ensing, C.G. de Leániz, E. García-Vázquez, A.M. Griffiths, K. Hindar, S. Hjorleifsdottir, D. Knox, G. Machado-Schiaffino, P. McGinnity, D. Meldrup, E.E. Nielsen, K. Olafsson, C.R. Primmer, P. Prodohl, L. Stradmeyer, J.P. Vähä, E. Verspoor, V. Wennevik, J.R. Stevens, Microsatellite standardization and evaluation of genotyping error in a large multi-partner research programme for conservation of Atlantic salmon (*Salmo salar*), Genetica 139 (2011) 353–367, http://dx.doi.org/10.1007/s10709-011-9554-4.
[29] L. Benestan, T. Gosselin, C. Perrier, B. Sainte-Marie, R. Rochette, L. Bernatchez, RAD-genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species; the American lobster (*Homarus americanus*), Mol. Ecol. 24 (2015) 3299–3315, http://dx.doi.org/10.1111/mec.13245.
[30] P. Marti-Puig, F. Costantini, L. Rugiu, M. Ponti, M. Abbiati, Patterns of genetic connectivity in invertebrates of temperate MPA networks, Adv. Oceanogr. Limnol. 4 (2013) 138–149, http://dx.doi.org/10.1080/19475721.2013.850445.
[31] M.A. Coleman, J. Chambers, N.A. Knott, H.A. Malcolm, D. Harasti, A. Jordan, B.P. Kelaher, Connectivity within and among a network of temperate marine reserves, PLoS One 6 (2011) e20168, http://dx.doi.org/10.1371/journal.pone.0020168.
[32] L.P. Holland, T.L. Jenkins, J.R. Stevens, Contrasting patterns of population structure and gene flow facilitate exploration of connectivity in two widely distributed temperate octocorals, Heredity 119 (2017) 35–48, http://dx.doi.org/10.1038/hdy.2017.14.
[33] J. Lubchenco, Ecology: the sea-otter whisperer, Nature 533 (2016) 318–319, http://dx.doi.org/10.1038/533318a.
[34] E. Pante, N. Puillandre, A. Viricel, S. Arnaud-Haond, D. Aurelle, M. Castelin, A. Chenuil, C. Destombe, D. Forcioli, M. Valero, F. Viard, S. Samadi, Species are hypotheses: avoid connectivity assessments based on pillars of sand, Mol. Ecol. 24 (2015) 525–544, http://dx.doi.org/10.1111/mec.13048.
[35] M.B.O. Huserbraten, E. Moland, H. Knutsen, E.M. Olsen, C. André, N.C. Stenseth, Conservation, spillover and gene flow within a network of northern European Marine Protected Areas, PLoS One 8 (2013) e73388, http://dx.doi.org/10.1371/journal.pone.0073388.
[36] B.J. Puckett, D.B. Eggleston, P.C. Kerr, R.A. Luettich, Larval dispersal and population connectivity among a network of marine reserves, Fish. Oceanogr. 23 (2014) 342–361, http://dx.doi.org/10.1111/fog.12067.
[37] Y. Li, W. Wang, X. Liu, W. Luo, J. Zhang, Y. Gul, DNA extraction from crayfish exoskeleton, Indian J. Exp. Biol. 49 (2011) 953–957.
[38] M. Miller, J. Dunham, A. Amores, W. Cresko, E. Johnson, Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD), Genome Res. 17 (2007) 240–248, http://dx.doi.org/10.1101/gr.5681207.
[39] C.F. Graham, T.C. Glenn, A.G. McArthur, D.R. Boreham, T. Kieran, S. Lance, R.G. Manzon, J.A. Martino, T. Pierson, S.M. Rogers, J.Y. Wilson, C.M. Somers, Impacts of degraded DNA on restriction enzyme associated DNA Sequencing (RADSeq), Mol. Ecol. Resour. 15 (2015) 1304–1315, http://dx.doi.org/10.1111/1755-0998.12404.
[40] N.A. Baird, P.D. Etter, T.S. Atwood, M.C. Currey, A.L. Shiver, Z.A. Lewis, E.U. Selker, W.A. Cresko, E.A. Johnson, Rapid SNP discovery and genetic mapping using sequenced RAD markers, PLoS One 3 (2008) e3376, http://dx.doi.org/10.1371/journal.pone.0003376.
[41] C. Schlötterer, The evolution of molecular markers - just a matter of fashion? Nat. Rev. Genet. 5 (2004) 63–69, http://dx.doi.org/10.1038/nrg1249.
[42] F.W. Allendorf, P.A. Hohenlohe, G. Luikart, Genomics and the future of conservation genetics, Nat. Rev. Genet. 11 (2010) 697–709, http://dx.doi.org/10.1038/nrg2844.
[43] J. Cuéllar-Pinzón, P. Presa, S.J. Hawkins, A. Pita, Genetic markers in marine fisheries: types, tasks and trends, Fish. Res. 173 (2016) 194–205, http://dx.doi.org/10.1016/j.fishres.2015.10.019.
[44] F.W. Allendorf, Genetics and the conservation of natural populations: allozymes to genomes, Mol. Ecol. 26 (2017) 420–430, http://dx.doi.org/10.1111/mec.13948.
[45] S. Hoban, O. Gaggiotti, G. Bertorelle, Sample planning optimization tool for conservation and population genetics (SPOTG): a software for choosing the appropriate number of markers and samples, Methods Ecol. Evol. 4 (2013) 299–303, http://dx.doi.org/10.1111/2041-210x.12025.
[46] M.H. Meek, M.R. Baerwald, M.R. Stephens, A. Goodbla, M.R. Miller, K.M.H. Tomalty, B. May, Sequencing improves our ability to study threatened migratory species: genetic population assignment in California's Central Valley Chinook salmon, Ecol. Evol. 6 (2016) 7706–7716, http://dx.doi.org/10.1002/ece3.2493.
[47] T.L. Jenkins, C.D. Ellis, J.R. Stevens, SNP discovery in European lobster (*Homarus gammarus*) using RAD sequencing, Conserv. Genet. Resour. (2018), http://dx.doi.

org/10.1007/s12686-018-1001-8.
[48] W. Jiao, X. Fu, J. Li, L. Li, L. Feng, J. Lv, L. Zhang, X. Wang, Y. Li, R. Hou, L. Zhang, X. Hu, S. Wang, Z. Bao, Large-scale development of gene-associated single-nucleotide polymorphism markers for molluscan population genomic, comparative genomic, and genome-wide association studies, DNA Res. 21 (2014) 183–193, http://dx.doi.org/10.1093/dnares/dst048.
[49] L. Laikre, M.K. Schwartz, R.S. Waples, N. Ryman, Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals, Trends Ecol. Evol. 25 (2010) 520–529, http://dx.doi.org/10.1016/j.tree.2010.06.013.
[50] A.B.A. Shafer, J.B.W. Wolf, P.C. Alves, L. Bergstrom, M.W. Bruford, I. Brannstrom, G. Colling, L. Dalen, L. De Meester, R. Ekblom, K.D. Fawcett, S. Fior, M. Hajibabaei, J.A. Hill, A.R. Hoezel, J. Hoglund, E.L. Jensen, J. Krause, T.N. Kristensen, M. Krutzen, J.K. McKay, A.J. Norman, R. Ogden, E.M. Osterling, N.J. Ouborg, J. Piccolo, D. Popovic, C.R. Primmer, F.A. Reed, M. Roumet, J. Salmona, T. Schenekar, M.K. Schwartz, G. Segelbacher, H. Senn, J. Thaulow, M. Valtonen, A. Veale, P. Vergeer, N. Vijay, C. Vila, M. Weissensteiner, L. Wennerstrom, C.W. Wheat, P. Zielinski, Genomics and the challenging translation into conservation practice, Trends Ecol. Evol. 30 (2015) 78–87, http://dx.doi.org/10.1016/j.tree.2014.11.009.
[51] I.J. Gordon, D.M. Evans, T.W.J. Garner, T. Katzner, M.E. Gompper, R. Altwegg, T.A. Branch, J.A. Johnson, N. Pettorelli, Enhancing communication between conservation biologists and conservation practitioners: letter from the Conservation Front Line, Anim. Conserv. 17 (2014) 1–2, http://dx.doi.org/10.1111/acv.12097.
[52] B.A. Garner, B.K. Hand, S.J. Amish, L. Bernatchez, J.T. Foster, K.M. Miller, P.A. Morin, S.R. Narum, S.J. O'Brien, G. Roffler, W.D. Templin, P. Sunnucks, J. Strait, K.I. Warheit, T.R. Seamons, J. Wenburg, J. Olsen, G. Luikart, Genomics in conservation: case studies and bridging the gap between data and application, Trends Ecol. Evol. 31 (2016) 81–82, http://dx.doi.org/10.1016/j.tree.2015.10.009.
[53] C.J. Hogg, C.E. Grueber, D. Pemberton, S. Fox, A.V. Lee, J.A. Ivy, K. Belov, "Devil Tools & Tech": a synergy of conservation research and management practice, Conserv. Lett. 10 (2017) 133–138, http://dx.doi.org/10.1111/conl.12221.
[54] S.J. Galla, T.R. Buckley, R. Elshire, M.L. Hale, M. Knapp, J. Mccallum, R. Moraga, A.W. Santure, P. Wilcox, T.E. Steeves, Building strong relationships between conservation genetics and primary industry leads to mutually beneficial genomic advances, Mol. Ecol. 25 (2016) 5267–5281, http://dx.doi.org/10.1111/mec.13837.
[55] H.R. Taylor, N. Dussex, Y. van Heezik, Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners, Glob. Ecol. Conserv. 10 (2017) 231–242, http://dx.doi.org/10.1016/j.gecco.2017.04.001.
[56] W.E. Johnson, D.P. Onorato, M.E. Roelke, E.D. Land, M. Cunningham, R.C. Belden, R. McBride, D. Jansen, M. Lotz, D. Shindle, J. Howard, D.E. Wildt, L.M. Penfold, J.A. Hostetler, M.K. Oli, S.J. O'Brien, Genetic restoration of the Florida panther, Science 329 (2010) 1641–1645, http://dx.doi.org/10.1126/science.1192891.
[57] F.G. Çilingir, F.E. Rheindt, K.M. Garg, K. Platt, S.G. Platt, D.P. Bickford, Conservation genomics of the endangered Burmese roofed turtle, Conserv. Biol. (2017), http://dx.doi.org/10.1111/cobi.12921.
[58] M. Baerwald, M. Stephens, K. Bork, M. Meek, K. Tomalty, B. May, Spring-run Chinook salmon genetic management plan - San Joaquin river restoration program, University of California, 2011.
[59] C. Habicht, A.R. Munro, T.H. Dann, D.M. Eggers, W.D. Templin, M.J. Witteveen, T. T. Baker, K.G. Howard, J.R. Jasper, S.D. Rogers Olive, H.L. Liller, E.L. Chenoweth, E.C. Volk, Harvest and harvest rates of sockeye salmon stocks in fisheries of the western Alaska salmon stock identification program (WASSIP), Alaska Department of Fish and Game, Special Publication No. 12-24, Anchorage, 2012.
[60] F. Fiorentino, E. Massutì, F. Tinti, S. Somarakis, G. Garofalo, T. Russo, M. Facchini, P. Carbonara, K. Kapiris, P. Tugores, R. Cannas, C. Tsigenopoulos, B. Patti, F. Colloca, M. Sbrana, R. Mifsud, V. Valavanis, M.T. Spedicato, Stock units: Identification of distinct biological units (stock units) for different fish and shellfish species and among different GFCM-GSA, STOCKMED Deliverable 03: FINAL REPORT, September 2014, 310 p, 2014.
[61] S.G. Vandamme, A.M. Griffiths, S.-A. Taylor, C. Di Muri, E.A. Hankard, J.A. Towne, M. Watson, S. Mariani, Sushi barcoding in the UK: another kettle of fish, PeerJ 4 (2016) e1891, http://dx.doi.org/10.7717/peerj.1891.
[62] J.T. Martinsohn, R. Ogden, FishPopTrace-developing SNP-based population genetic assignment methods to investigate illegal fishing, Forensic Sci. Int. Genet. Suppl. Ser. 2 (2009) 294–296, http://dx.doi.org/10.1016/j.fsigss.2009.08.108.
[63] L.M. Lieberknecht, P.J.S. Jones, From stormy seas to the doldrums: the challenges of navigating towards an ecologically coherent marine protected area network through England's Marine Conservation Zone process, Mar. Policy 71 (2016) 275–284, http://dx.doi.org/10.1016/j.marpol.2016.05.023.
[64] L.M. Lieberknecht, T.W. Mullier, J.A. Ardron, Assessment of the ecological coherence of the UK's marine protected area network. A report prepared for the Joint Links, 2014.
[65] H. Carr, H. Wright, A. Cornthwaite, J. Davies, Assessing the contribution of Welsh MPAs towards an ecologically coherent MPA network in 2016, JNCC Report. 〈http://jncc.defra.gov.uk/pdf/JNCC_NetworkProgressWelshWaters_Final.pdf〉.
[66] C.M. Roberts, J.P. Hawkins, J. Fletcher, S. Hands, K. Raab, S. Ward, Guidance on the size and spacing of Marine Protected Areas in England, Natural England, Commissioned Report NECR037. 〈http://publications.naturalengland.org.uk/publication/46009〉.
[67] A.S. Kough, C.B. Paris, The influence of spawning periodicity on population connectivity, Coral Reefs 34 (2015) 753–757, http://dx.doi.org/10.1007/s00338-015-1311-1.
[68] F. Palero, P. Abelló, E. Macpherson, M. Gristina, M. Pascual, Phylogeography of the European spiny lobster (*Palinurus elephas*): influence of current oceanographical features and historical processes, Mol. Phylogenet. Evol. 48 (2008) 708–717, http://dx.doi.org/10.1016/j.ympev.2008.04.022.
[69] J.K. Pritchard, M. Stephens, P. Donnelly, Inference of population structure using multilocus genotype data, Genetics 155 (2000) 945–959, http://dx.doi.org/10.1111/j.1471-8286.2007.01758.x.
[70] S.K. Pikesley, B.J. Godley, H. Latham, P.B. Richardson, L.M. Robson, J.-L. Solandt, C. Trundle, C. Wood, M.J. Witt, Pink sea fans (*Eunicella verrucosa*) as indicators of the spatial efficacy of Marine Protected Areas in southwest UK coastal waters, Mar. Policy 64 (2016) 38–45, http://dx.doi.org/10.1016/j.marpol.2015.10.010.
[71] R.K. Cowen, G. Gawarkiewicz, J. Pineda, S.R. Thorrold, F.E. Werner, Population connectivity in marine systems, Oceanography 20 (2007) 14–21, http://dx.doi.org/10.1126/science.1122039.
[72] C. Riginos, E.D. Crandall, L. Liggins, P. Bongaerts, E.A. Treml, Navigating the currents of seascape genomics: how spatial analyses can augment population genomic studies, Curr. Zool. 62 (2016) 581–601, http://dx.doi.org/10.1093/cz/zow067.
[73] K.A. Selkoe, C.C. D'Aloia, E.D. Crandall, M. Iacchei, L. Liggins, J.B. Puritz, S. Von Der Heyden, R.J. Toonen, A decade of seascape genetics: contributions to basic and applied marine connectivity, Mar. Ecol. Prog. Ser. 554 (2016) 1–19, http://dx.doi.org/10.3354/meps11792.
[74] C. Breusing, A. Biastoch, A. Drews, A. Metaxas, D. Jollivet, R.C. Vrijenhoek, T. Bayer, F. Melzner, L. Sayavedra, J.M. Petersen, N. Dubilier, M.B. Schilhabel, P. Rosenstiel, T.B.H. Reusch, Biophysical and population genetic models predict the presence of "phantom" stepping stones connecting Mid-Atlantic ridge vent ecosystems, Curr. Biol. 26 (2016) 2257–2267, http://dx.doi.org/10.1016/j.cub.2016.06.062.
[75] D. Hedgecock, P.H. Barber, S. Edmands, Genetic approaches to measuring connectivity, Oceanography 20 (2007) 70–79.

OPEN

Heredity (2017) 119, 35–48
Official journal of the Genetics Society
www.nature.com/hdy

ORIGINAL ARTICLE

Contrasting patterns of population structure and gene flow facilitate exploration of connectivity in two widely distributed temperate octocorals

LP Holland[1], TL Jenkins[1] and JR Stevens

Connectivity is an important component of metapopulation dynamics in marine systems and can influence population persistence, migration rates and conservation decisions associated with Marine Protected Areas (MPAs). In this study, we compared the genetic diversity, gene flow and population structure of two octocoral species, *Eunicella verrucosa* and *Alcyonium digitatum,* in the northeast Atlantic (ranging from the northwest of Ireland and the southern North Sea, to southern Portugal), using two panels of 13 and 8 microsatellite loci, respectively. Our results identified regional genetic structure in *E. verrucosa* partitioned between populations from southern Portugal, northwest Ireland and Britain/France; subsequent hierarchical analysis of population structure also indicated reduced gene flow between southwest Britain and northwest France. However, over a similar geographical area, *A. digitatum* showed little evidence of population structure, suggesting high gene flow and/or a large effective population size; indeed, the only significant genetic differentiation detected in *A. digitatum* occurred between North Sea samples and those from the English Channel/northeast Atlantic. In both species the vast majority of gene flow originated from sample sites within regions, with populations in southwest Britain being the predominant source of contemporary exogenous genetic variants for the populations studied. Overall, historical patterns of gene flow appeared more complex, though again southwest Britain appeared to be an important source of genetic variation for both species. Our findings have major conservation implications, particularly for *E. verrucosa*, a protected species in UK waters and listed by the IUCN as 'Vulnerable', and for the designation and management of European MPAs.
Heredity (2017) **119,** 35–48; doi:10.1038/hdy.2017.14; published online 15 March 2017

INTRODUCTION

Population connectivity has emerged as a key factor in the sustainable management of marine resources (Fogarty and Botsford, 2007; Da Silva *et al.*, 2014), in tracking invasive species (Pérez-Portela *et al.*, 2012), in monitoring the effects of climate change (Munday *et al.*, 2009; Gerber *et al.*, 2014), and in designating networks of protected areas (Jones *et al.*, 2007; Marti-Puig *et al.*, 2013). For most benthic marine organisms, connectivity is typically defined by dispersal during early life stages and is intimately associated with oceanic currents and topographical features (Cowen *et al.*, 2007). However, connectivity can vary across marine taxa, even between closely related species over similar spatial scales (Bargelloni *et al.*, 2003, 2005; Charrier *et al.*, 2006; Kool *et al.*, 2013) and population structure can be determined by the extent of dispersal from distant vs local sources, resulting in fully 'open' (panmictic) to fully 'closed' (isolated) populations (see Cowen and Sponaugle, 2009 and references therein). Perhaps most importantly from an applied perspective, population structure and gene flow can be used as a proxy for understanding population connectivity (Hedgecock *et al.*, 2007; Lowe and Allendorf, 2010; Kool *et al.*, 2013).

Advances in our knowledge of marine population connectivity are fundamental for the strategic allocation of available resources in a way that maximises protection of marine biodiversity (Kool *et al.*, 2013). Moreover, the global extent of protected areas is unlikely to mitigate the current rate of marine and terrestrial biodiversity loss (Mora and Sale, 2011). Among the 15 European countries that have signed the Oslo/Paris (OSPAR) Convention (for the protection of the marine environment of the northeast Atlantic), there is a requirement to establish an 'ecologically coherent' network of Marine Protected Areas (MPAs), which collectively aims to deliver more benefits to biodiversity than single, unrelated MPAs (OSPAR Convention, 2013). As connectivity is a key feature of an MPA network, it is important that empirical estimates of population connectivity are considered during the designation or review stages of a network (Jones *et al.*, 2007). For example, guidelines for incorporating connectivity into designing networks of marine reserves are available for coral reefs and are likely to be useful for the management and protection of these systems (Almany *et al.*, 2009; McCook *et al.*, 2009). Several analyses of connectivity in established networks have also identified deficiencies that may reduce the efficacy of a network. For example, Puckett *et al.* (2014) modelled dispersal of the eastern oyster (*Crassostrea virginica*) on the Atlantic coast of North Carolina and showed that if marine reserves were too small – relative to the mean dispersal distance of the oyster– local retention of larvae was reduced; likewise, if reserves were spaced too far apart, connectivity became limited. While early

Department of Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter, UK
Correspondence: Dr J Stevens, Department of Biosciences, University of Exeter, Geoffrey Pope Building, Stocker Road, Exeter EX4 4QD, UK.
E-mail: j.r.stevens@exeter.ac.uk
[1]These authors contributed equally to this work.
Received 24 May 2016; revised 3 February 2017; accepted 17 February 2017; published online 15 March 2017

landmark studies of genetic connectivity in the marine environment (for example, Palumbi, 2003) focused largely on gene flow, barriers to gene flow and isolation by distance (IBD), more recent studies (for example, Arizmendi-Mejía *et al.*, 2015; Gagnaire *et al.*, 2015) have further refined our understanding of drivers of marine genetic connectivity and have demonstrated the importance of additional factors in driving or disrupting genetic connectivity, for example, effective population size and genetic drift. Overall, such findings suggest a greater understanding of population structure and connectivity is required to optimise the conservation of marine biodiversity and to maximise the efficacy of such networks (for example, OSPAR Commission, 2006; Jones *et al.*, 2007).

Currently, and until such time as a robust understanding of the functioning of networks of MPAs is achieved, individual MPAs are typically designated based on the presence of rare or protected species or guilds of species; for example, in the waters of southwest Britain, the presence of *Eunicella verrucosa* (the pink sea fan) is often listed as a factor in the designation of an area as a Marine Conservation Zone (MCZ). However, population genetic studies of octocorals across this area, and the northeast Atlantic in general, are limited. Previous research in this region has had either a phylogenetic (for example, McFadden and Hutchinson, 2004) or phylogeographic focus (for example, Herrera *et al.*, 2012), while existing connectivity research on this subclass in the region has assessed the genetic diversity and structure of primarily Mediterranean species, for example, *Corallium rubrum* (Costantini *et al.*, 2007; Ledoux *et al.*, 2010; Aurelle *et al.*, 2011), *Eunicella singularis* (Costantini *et al.*, 2016), *Eunicella cavolini* (Masmoudi *et al.*, 2016) and *Paramuricea clavata* (Mokhtar-Jamai *et al.*, 2011; Arizmendi-Mejía *et al.*, 2015). As a result, genetic diversity and connectivity in this group remains understudied.

Eunicella verrucosa is an IUCN red-listed octocoral. It can be found from Angola to western Ireland, but its range in the British Isles is limited to southwest England, southwest Wales, and southern and western Ireland (Hayward and Ryland, 1995). In Britain, it is considered rare due to its limited distribution beyond the southwest (Hiscock *et al.*, 2010), although where it is found it can be relatively abundant and may form 'forests'. Colonies are generally found inhabiting rocky substrates, at depths of 10–150 m, in areas of high turbidity with moderate to high current flow. *E. verrucosa* has an important role for the functional ecology of sublittoral ecosystems in which it occurs; it provides structural complexity and habitat for numerous epifauna and, as such, may be considered to be an ecosystem engineer (Hall-Spencer *et al.*, 2007; Pikesley *et al.*, 2016). Colonies are also vulnerable to trawling activity and, as a result, the designation of several MPAs across Britain includes *E. verrucosa* as a specific factor (a 'protected feature') in their designation.

Alcyonium digitatum ('dead man's fingers') has a ubiquitous presence along rocky upper and circalittoral zones, typically to a depth of 200 m, and it can be found around most British and Irish coasts (Hayward and Ryland, 1995); it is represented in several MPAs across the UK network. It is widely distributed across the North Atlantic, ranging from Portugal to Norway, to eastern Canada, south to Cape Hatteras in the USA (Hartnoll, 1975; Watling and Auster, 2005). It is not a protected species, however, it is locally depleted in some areas by benthic trawling (Hinz *et al.*, 2011). Both species are thought to be lecithotrophic, gonochoristic (separate sexes) and broadcast spawners, with limited reports of hermaphroditism in *A. digitatum*; asexual reproduction may also be possible in *E. verrucosa* as genets can proliferate via fragmentation (Hartnoll, 1975; McFadden *et al.*, 2001; Munro, 2004). *Alcyonium digitatum* releases gametes in winter (December–January) and pelagic larvae can survive up to 14 weeks and beyond (Hartnoll, 1975). Spawning of *E. verrucosa* occurs towards the end of summer (August–September), though its pelagic larval duration is unknown (Munro, 2004). Studying patterns in genetic connectivity and assessing genetic diversity offers an alternative approach by which to infer the dispersal capabilities of these species.

In this study, two panels of microsatellites (Holland *et al.*, 2013a, b) were used to assess the population structure and genetic connectivity of *E. verrucosa* and *A. digitatum* around the British Isles and northeast Atlantic. Specifically, we addressed the following questions: (i) what is the genetic diversity of each species and is it uniform across the sampling range; (ii) do both of these species show population genetic structure indicative of departures from panmixia; and (iii) what are the levels of gene flow and effective population size for each species? Finally, we consider the conservation and potential management implications of our findings for these species, both in terms of connectivity between existing MPAs and with regard to the designation of future MPAs.

MATERIALS AND METHODS

Study sites and sampling

Samples of *E. verrucosa* ($N = 922$) were collected from 27 sites ranging from southern Portugal to northwest Ireland, including sites around Brittany in northwest France, Lyme Bay in southern England and southwest Wales (Table 1 and Figure 1). The area sampled represents much of the northern range of the species. Samples of *A. digitatum* ($N = 655$) were collected from 20 sites across a similar geographic area (with the exception of southern Portugal, where the species was not found); samples from two additional sites in the North Sea (Table 2 and Figure 1) were also included. The area sampled represents much of the southern range of *A. digitatum* in Europe. Samples of both species were collected between 2007 and 2012. The majority of samples were collected by SCUBA at depths between 10 and 35 m; additional samples of *A. digitatum* were collected by trawling (CEFAS scientific trawl, Lowestoft, UK). Colonies of *E. verrucosa* were sampled by removing a 3 cm terminal branch using sea-snips. This species is protected in UK waters, and all UK sampling complied with licenses granted by Natural England and the Marine Management Organisation (see Acknowledgements). Colonies of *A. digitatum* were sampled by removing a 1 cm^3 section of tissue from a terminal thumb-like 'branch' using sea-snips. After removal, individual colonies were placed into mesh bags, brought to the surface, and quickly immersed in 95–100% ethanol for storage. In both species, samples were taken from individual colonies spaced at least 1 m apart to avoid sampling clones; previous studies of hard corals have identified potential clones at spatial scales up to 5 m apart (for example, Goffredo *et al.*, 2009; Foster *et al.*, 2012). This issue was also addressed after genotyping by identifying and excluding any duplicate genotypes occurring in the same population.

DNA extraction and microsatellite genotyping

Total genomic DNA was extracted from ~10 to 20 polyps using a WizardR SV Genomics DNA Purification System kit (Promega, Southampton, UK) following the manufacturer's protocol. Polyps were removed from colonies using forceps, or by using a scalpel to shave a portion of ~1 cm^2 surface tissue from *A. digitatum* or 1–2 cm of coenenchymal tissue (excluding the gorgonin axis) from *E. verrucosa*. Microsatellites were amplified for both species and alleles were scored using GENEMAPPER v3.7 (Applied Biosystems, Paisley, UK). Full details of DNA extraction and microsatellite amplification conditions and multiplexing are given in a primer note for each species: *E. verrucosa* (Holland *et al.*, 2013a) and *A. digitatum* (Holland *et al.*, 2013b).

Data screening and quality assessment

Duplicate genotypes were identified in CERVUS v3.0.3 (Kalinowski *et al.*, 2007) and were removed from further analyses. The presence of possible null alleles, allele scoring errors due to stuttering and large allele dropout was evaluated using MICRO-CHECKER v2.2.3 (Van Oosterhout *et al.*, 2004). Linkage

Table 1 Sampling information and summary statistics for *Eunicella verrucosa* samples

Region/Population	*Code*	*N*	N_g	*Depth (m)*	*Lat*	*Long*	H_{exp}	A_r	PA_r	F_{IS}
Britain										
[a]Isles of Scilly, Flat Ledge	Fla	23	23	30	49.97	−6.26	0.392	2.45	0.017	−0.021
[a]Isles of Scilly, Lion Rock	Lio	22	22	24	49.98	−6.31	0.435	2.66	0.017	0.016
[a]Lundy Island	Lun	23 (1)	22	23	51.17	−4.69	0.428	2.61	0.038	0.032
[b]Lyme Bay, The Heroine Wreck	Her	9	9	25	50.68	−2.94	0.432	2.79	0.044	0.006
[b]Lyme Bay, Sawtooth Ledges	Saw	12	12	22	50.68	−2.80	0.383	2.43	0.023	0.106
[b]Lyme Bay, West Tennents Reef	Wte	45 (2)	43	23	50.65	−2.96	0.452	2.74	0.028	0.052
[a]Manacles, Raglan Rocks	Rag	44 (1)	43	28	50.04	−5.04	0.438	2.64	0.036	0.017
[a]Manacles, SS Mohegan Wreck	Moh	30	30	26	50.05	−5.04	0.409	2.52	0.015	**0.140**
Porthallow Bay, Volnay Wreck	Vol	24	24	21	50.07	−5.00	0.401	2.51	0.023	0.002
[a]Padstow, Camel Estuary	Cam	11 (3)	7	n/a	50.59	−4.95	0.433	2.71	0.001	**−0.202**
Plymouth, Bovisand	Bov	40	40	10	50.34	−4.13	0.423	2.54	0.020	**0.086**
Plymouth, Hand Deeps	Han	36	36	25	50.21	−4.34	0.459	2.76	0.018	0.062
Plymouth, Mewstone	Mew	45 (1)	44	24	50.30	−4.11	0.451	2.65	0.027	0.045
[a]Skomer Island	Sko	39	39	22	51.74	−5.30	0.445	2.61	0.013	−0.013
Ireland										
Donegal, Black Rock	Bla	29	29	25	54.58	−8.43	0.367	2.38	0.034	−0.013
Sligo, Thumb Rock	Thu	48	48	20	54.47	−8.44	0.376	2.46	0.071	**0.097**
France										
Brittany, Rade de Brest	Bre	43	43	35	48.31	−4.42	0.412	2.55	0.026	0.047
Brittany, Laonegued Taer	Lao	40	40	30	47.73	−4.06	0.419	2.56	0.059	**0.082**
Brittany, Men Goe	Men	43	43	30	47.69	−3.99	0.418	2.54	0.035	0.055
Brittany, Roscoff1	Ros1	40	40	35	48.75	−3.96	0.419	2.56	0.024	0.014
Brittany, Roscoff2	Ros2	39 (3)	36	25	48.71	−3.90	0.448	2.68	0.049	**0.071**
Portugal										
Algarve, Portimao1	Por1	42	42	17	37.10	−8.58	0.429	2.63	0.038	**0.128**
Algarve, Portimao2	Por2	36 (1)	35	18	37.10	−8.56	0.435	2.63	0.058	**0.105**
Algarve, Armacao de Pera1	Arm1	27	27	28	37.09	−8.35	0.402	2.53	0.041	**0.096**
Algarve, Armacao de Pera2	Arm2	44 (1)	43	21	37.05	−8.35	0.392	2.47	0.025	0.028
Algarve, Armacao de Pera3	Arm3	44 (3)	41	25	37.04	−8.36	0.406	2.52	0.039	0.067
Algarve, Faro	Far	44	44	17	36.98	−7.99	0.402	2.53	0.033	**0.091**

Abbreviation: CI, confidence interval.
Number of individuals genotyped per population (N) (with number of duplicate genotypes), number of unique genotypes per population (N_g), expected heterozygosity (H_{exp}), allelic richness (A_r), private allelic richness (PA_r) and the inbreeding coefficient (F_{IS}) are reported for each population. F_{IS} values significantly different from zero (95% CI) are highlighted in bold.
[a]Marine Conservation Zone.
[b]Candidate Special Area of Conservation.

disequilibrium and deviation from Hardy–Weinberg equilibrium (HWE) were tested in GENEPOP v4.2 (Rousset, 2008) using default parameters and the false discovery rate was used to detect type-1 errors (Storey and Tibshirani, 2003). To identify candidate markers under selection or linked with markers under selection, loci were screened using two different F_{ST} outlier detection methods in Lositan (Antao *et al.*, 2008) and Arlequin v3.5.2 (Excoffier and Lischer, 2010). Lositan assumes an island model and runs were conducted using the infinite alleles model. Parameters were set to 50 000 simulations, a 99% confidence interval and a false discovery rate of 0.1, with the neutral and forced mean F_{ST} enforced. In Arlequin, 50 000 simulations were run with 100 demes simulated per group and 10 simulated groups under the hierarchical island model. Samples were grouped by geographical region (*E. verrucosa*: Portugal, France, Ireland, Britain; *A. digitatum*: France, Ireland, Britain, North Sea) and results were considered significant if the P-value was <0.010.

Genetic variation

Expected heterozygosity (H_{exp}) and the inbreeding coefficient (F_{IS}) for each population were estimated using the *diveRsity* package (Kennan *et al.*, 2013) in R (R Development Core Team, 2016). The *divBasic* function was used and F_{IS} significance was assessed using 95% confidence intervals using 1000 bootstrap replicates; the significance level for multiple comparisons was corrected using a Bonferroni correction (Dunn, 1961), which had the effect of slightly widening each interval. Allelic richness (A_r) and private allelic richness (PA_r) were calculated in HP-RARE v1.1 (Kalinowski, 2005) using a rarefaction method, which accounts for variation in sample size (each sample included a minimum of eight loci).

Population structure

Population differentiation was analysed using pairwise F_{ST} (Weir and Cockerham, 1984) and G''_{ST} (Meirmans and Hedrick, 2011) measures using the *diffCalc* function in *diveRsity*, and significance was assessed as for F_{IS}. To search for genetic structuring, a principal coordinates analysis (PCoA) was performed using a matrix of codominant genotypic genetic distances in GenALEx v6.5 (Peakall and Smouse, 2012). An analysis of molecular variance (AMOVA) was performed using Arlequin (10 000 permutations) to test for differentiation amongst geographical regions. Population structure was also analysed using a Bayesian clustering method: STRUCTURE v2.3.4 (Pritchard *et al.*, 2000), using a burn-in of 10^4 and 10^6 repetitions. An admixture ancestry model using population IDs as priors and correlated allele frequencies was chosen. To determine the number of populations (K), the delta K statistic (Evanno *et al.*, 2005) and the mean value of L(K) were examined in the POPHELPER R package (Francis, 2017). Ten replicate runs were aligned and merged in POPHELPER using

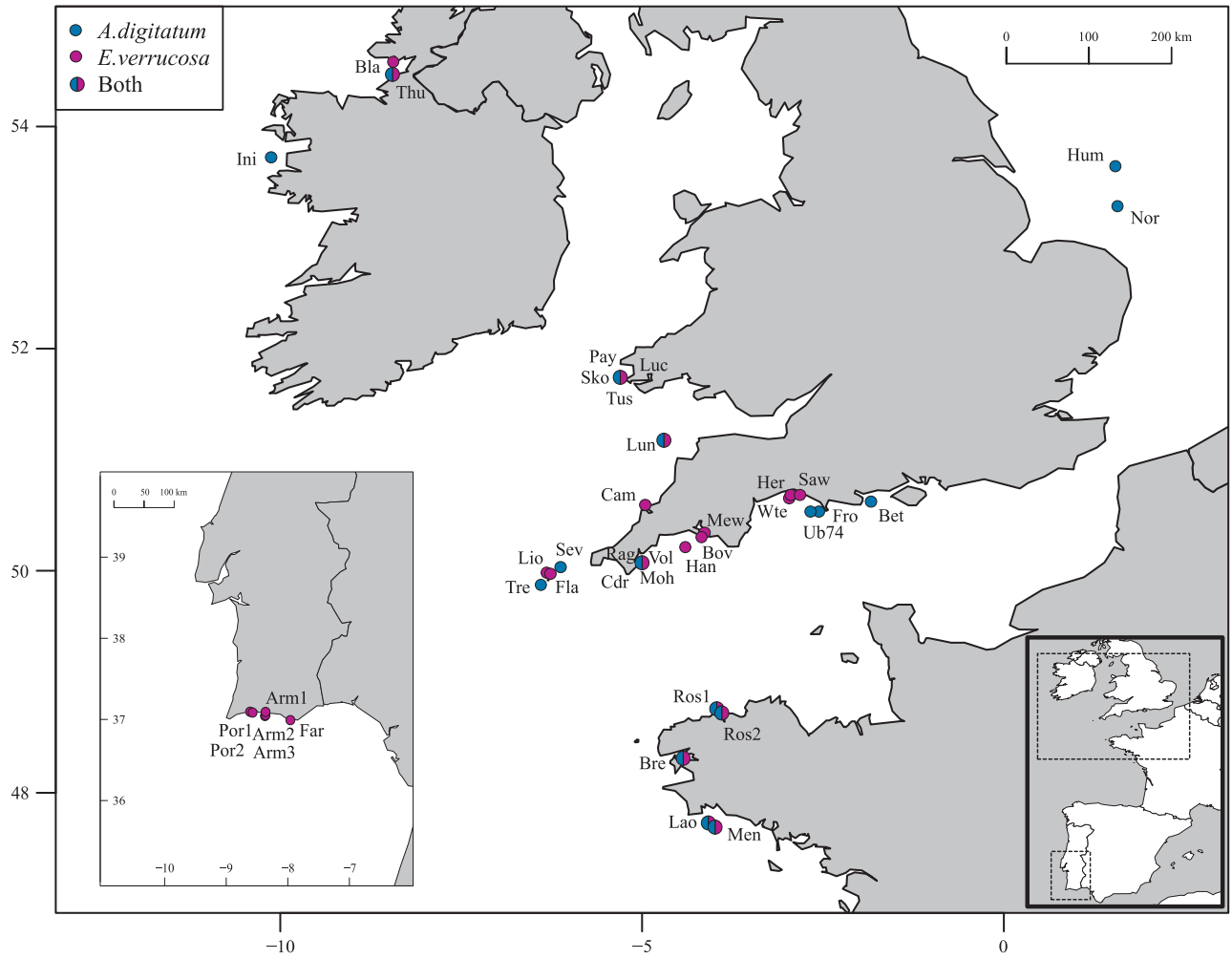


Figure 1 Map of the sites sampled in the northeast Atlantic. Pink circles represent sites where only *Eunicella verrucosa* were collected and blue circles represent where only *Alcyonium digitatum* were collected. Circles containing both colours represent sites in which both *E. verrucosa* and *A. digitatum* were collected. See Table 1 for details on population codes, sample size and latitude and longitude.

CLUMPP and graphics were generated using the merged data. Initial runs for both species using K values 1–10 showed a very low likelihood for K values 6–10, therefore, subsequent runs included only K values of 1–5. A Mantel test was implemented in GenALEx to test whether any observed genetic structure was a product of IBD. Genetic distances were supplied as $F_{ST}/(1- F_{ST})$ matrices and were compared with the logarithm of geographic distances (km). Negative F_{ST} values were converted to zero for this analysis. Geographical distances were estimated in Google Earth by measuring the shortest in-water distance between sites in a straight line or by calculating the shortest distance following coastlines.

Gene flow and effective population size

Contemporary (within the last few generations) and historical gene flow was estimated using two methods. Contemporary gene flow was analysed using BayesAss v3.0.4 (Wilson and Rannala, 2003), which estimates the fraction of immigrants in a population using Bayesian inference. Three runs were performed using 10^7 iterations, a burn-in of 10^6 and a sampling interval of 100, and an average of the gene flow estimates was calculated. The mixing parameters DeltaA, DeltaF and DeltaM were set to 0.10, 0.20 and 0.05 for *E. verrucosa*, and 0.30, 0.50 and 0.10 for *A. digitatum*, respectively. Convergence of the chains was validated using Tracer v1.6 (Rambaut *et al.*, 2014). Historical gene flow was calculated using the mutation-scaled migration rate M (m/μ; where m is the immigration rate per generation) and the population parameter theta ($4N_e{}^*\mu$) in Migrate-n v3.6 (Beerli and Felsenstein, 2001). Migrate-n is a coalescence-based program that has the benefit of providing values of immigration and emigration for each population and is therefore useful in scenarios of asymmetrical migration. A Brownian motion model was used and assumed a migration matrix with variable theta and estimated mutation rates for loci based on the data. A Bayesian likelihood strategy was initially run with default parameters to obtain start parameter estimates for theta and M. These parameters were supplied to the program in subsequent runs and the number of recorded steps was increased to 50 000. Prior uniform distributions for theta and M were set to min = 0, max = 100 and delta = 10, and min = 0, max = 1000 and delta = 100, respectively. To evaluate convergence of the chains, the effective sample size (>1000) and the shape of the histograms in the output files were examined.

Migrate-n was also used to calculate the mutation-scaled effective population size (N_e). This was calculated from the optimum value of theta using the equation $N_e = \text{theta}/4\,\mu$, assuming a microsatellite mutation rate (μ) of 10^{-4} per generation, as used in a previous study of a Mediterranean cup coral (Casado-Amezúa *et al.*, 2012). Contemporary N_e was estimated using LDNE v1.31 (Waples and Do, 2008). The program was run assuming a model of random mating and the allowed frequency of observed alleles was set to 0.050.

RESULTS

Data screening and quality assessment

For *E. verrucosa*, based on evidence of null alleles and significant deviation from HWE, one locus (Ever009) was omitted from the original microsatellite panel of Holland *et al.* (2013a). Five other loci

Table 2 Sampling information and summary statistics for *Alcyonium digitatum* samples

Region/Population	*Code*	*N*	N_g	*Depth (m)*	*Lat*	*Long*	H_{exp}	A_r	PA_r	F_{IS}
Britain										
[a]Isles of Scilly, Seven Stones Reef	Sev	40	40	35	50.03	−6.12	0.624	4.12	0.162	**0.055**
[a]Isles of Scilly, Trenemene Reef	Tre	42	42	32	49.87	−6.39	0.598	4.07	0.133	0.003
[a]Lundy Island	Lun	36	36	23	51.20	−4.68	0.618	4.15	0.098	0.023
[b]Lyme Bay, Frognor Wreck	Fro	18	18	34	50.53	−2.55	0.626	4.25	0.193	0.042
[b]Lyme Bay, UB74 Wreck	Ub74	19	19	34	50.53	−2.56	0.635	4.18	0.130	0.034
[a]Manacles, Carn-du-rocks	Cdr	35 (2)	33	26	50.05	−5.05	0.636	4.21	0.122	0.007
Porthallow Bay, Volnay Wreck	Vol	28	28	21	50.07	−5.00	0.659	4.31	0.099	−0.001
[a]Skomer Island, The Lucy Wreck	Luc	23 (1)	22	35	51.74	−5.28	0.594	4.01	0.075	0.027
[a]Skomer Island, Payne's Rock	Pay	51	51	30	51.74	−5.31	0.637	4.17	0.109	0.025
[a]Skomer Island, Tusker Rock	Tus	21	21	29	51.74	−5.26	0.629	4.22	0.189	0.031
Swanage, Betsy Anna Wreck	Bet	26 (2)	24	23	50.62	−1.83	0.620	4.07	0.085	0.024
North Sea, Humberside	Hum	27	27	25	53.64	1.55	0.618	4.08	0.120	−0.014
North Sea, Norfolk	Nor	33	33	25	53.28	1.58	0.668	4.29	0.106	0.059
Ireland										
Mayo, Inishturk Island	Ini	48	48	27	53.72	−10.12	0.625	4.13	0.118	0.038
Sligo, Thumb Rock	Thu	18	18	15	54.47	−8.44	0.655	4.24	0.156	−0.058
France										
Brittany, Rade de Brest	Bre	43	43	35	48.34	−4.58	0.645	4.23	0.134	**0.068**
Brittany, Laonegued Taer	Lao	29	29	30	47.73	−4.06	0.595	3.99	0.133	0.053
Brittany, Men Goe	Men	35 (1)	34	30	47.69	−3.99	0.653	4.22	0.087	**0.063**
Brittany, Roscoff1	Ros1	41	41	35	48.75	−3.96	0.635	4.23	0.199	0.035
Brittany, Roscoff2	Ros2	42 (1)	41	25	48.71	−3.90	0.649	4.34	0.117	**0.062**

Abbreviation: CI, confidence interval.
Number of individuals genotyped per population (*N*) (with number of duplicate genotypes), number of unique genotypes per population (N_g), expected heterozygosity (H_{exp}), allelic richness (A_r), private allelic richness (PA_r) and the inbreeding coefficient (F_{IS}) are reported for each population. F_{IS} values significantly different from zero (95% CI) are highlighted in bold.
[a]Marine Conservation Zone.
[b]Candidate Special Area of Conservation.

also showed some deviation from HWE, however, departures from HWE occurred in only a few populations and these loci were retained. Similarly, linkage disequilibrium was detected in five populations, but each population showed different pairs of potentially linked loci. With no obvious trend in the pattern of linkage disequilibrium observed, this inconsistency was likely due to site-specific biological processes which we were not able to investigate further within this study; consequently, no loci were discarded on the basis of linkage disequilibrium and 13 were used for subsequent analyses. For *A. digitatum*, three loci (Adig003, Adig004 and Adig010) were discarded from the original microsatellite panel of Holland *et al.* (2013b) based on the presence of null alleles and significant deviations from HWE. Some evidence of linkage disequilibrium was also detected, but was minimal across populations and no further loci were omitted; eight loci were used for subsequent analyses.

A relatively low number of duplicate genotypes were identified in both species. In *E. verrucosa*, 17 individuals with duplicate genotypes were identified in nine samples (Table 1), while in *A. digitatum*, seven individuals with duplicate genotypes were identified in five samples (Table 2). Duplicates were removed from further analyses. The spread of duplicates across sites did not show any obvious pattern in either species, with the exception of a small sample of *E. verrucosa* (Cam) from north Cornwall, in which four duplicate individuals (across three genotypes) were identified out of a sample of only 11 individuals successfully genotyped.

E. verrucosa samples were monomorphic at several loci, but this was not consistent in all populations at the same locus. In comparison, *A. digitatum* was monomorphic at only one locus (Adig007) in three populations. For *E. verrucosa*, two loci (Ever013 and Ever014) were identified as outliers under the island model and one (Ever013) under the hierarchical island model (Supplementary Appendix 1). Accordingly, as both methods identified Ever013 as an outlier under positive selection, analyses of population structure, gene flow and effective population size excluded this locus; STRUCTURE, PCoA and BayesAss analyses were conducted using 13 loci as the assumptions of these methods are not violated by the inclusion of loci under selection (Pritchard *et al.*, 2000; Wilson and Rannala, 2003). One outlier locus (Adig006) was identified for *A. digitatum* by the island model, but not by the hierarchical island model (Supplementary Appendix 1); accordingly, eight loci were retained.

Genetic variation

After removal of duplicate genotypes, genotypes of 905 individual specimens of *E. verrucosa* from 27 sites were analysed at 13 loci. For *A. digitatum*, genotypes of 648 individual specimens from 20 sites were analysed at eight loci. For *E. verrucosa*, measures of H_{exp} ranged from 0.367 (Black Rock) to 0.459 (Hand Deeps) and were generally consistent within regions, with minor differences between some regions (Table 1). A similar pattern was observed for A_r, which ranged from 2.38 (Black Rock) to 2.79 (The Heroine Wreck); overall, both measures were slightly lower in the samples from Ireland. For *A. digitatum*, H_{exp} and A_r measures were also relatively uniform within and between regions (Table 2) and were consistently higher than for *E. verrucosa*; H_{exp} measures ranged from 0.594 (The Lucy Wreck) to 0.668 (Norfolk) and A_r ranged from 3.99 (Laonegued Taer) to 4.34 (Roscoff2). Private allelic richness (PA_r) was also consistently higher

for *A. digitatum* than for *E. verrucosa*; values for *A. digitatum* ranged from 0.075 (The Lucy Wreck) to 0.193 (Roscoff1), while values for *E. verrucosa* were between 0.001 (Camel Estuary) to 0.059 (Laonegued Taer). The majority of F_{IS} values for both species were positive; overall, however, few were significant, though generally at least one site in each region showed a significant positive F_{IS} coefficient (Tables 1 and 2). This finding suggested a deficiency of heterozygotes at some sites; for *E. verrucosa*, this was most apparent in several populations from Portugal, while both species showed significant, positive F_{IS} values at Roscoff2. The broader implications of these findings (inbreeding and/or a Wahlund effect caused by the inadvertent combining of data from separate populations) are discussed below. A small sample of *E. verrucosa* from the Camel Estuary (Cam) had a significantly negative F_{IS}, indicating an excess of heterozygotes at this site.

Population structure

Global F_{ST} and G''_{ST} measures across all populations of *E. verrucosa* were 0.012 and 0.023, respectively (Supplementary Appendix 2). In comparison, global F_{ST} and G''_{ST} measures for all populations of *A. digitatum* were lower (0.003 and 0.011, respectively) (Supplementary Appendix 2). For both species, global values were significantly different from zero.

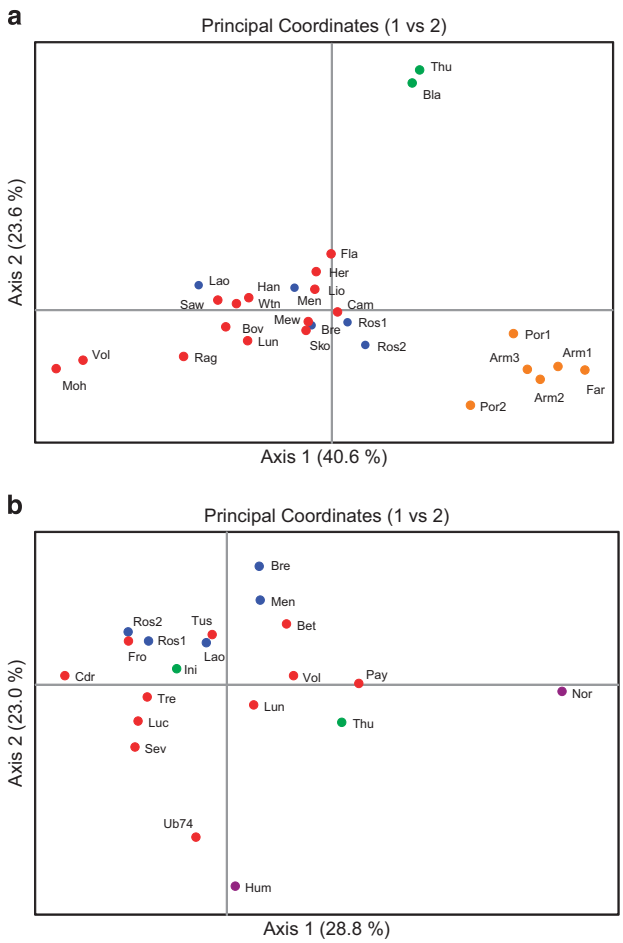


Figure 2 PCoA for *Eunicella verrucosa* (**a**) and *Alcyonium digitatum* (**b**). Colours correspond to regions: Britain (red), France (blue), Ireland (green), Portugal (orange), the North Sea (purple).

For *E. verrucosa*, the largest significant pairwise F_{ST} (0.059) value was observed between Faro and the Camel Estuary, while the highest significant pairwise G''_{ST} (0.089) value observed was also between populations from Portugal and southwest Britain: Portamao2 and the Heroine Wreck (Supplementary Appendix 2). In contrast, the highest significant pairwise F_{ST} and G''_{ST} values (0.020 and 0.058, respectively) for *A. digitatum* were between populations from southwest Britain and the North Sea: Trenemene Reef and Norfolk (Supplementary Appendix 2). For both species, both pairwise measures were typically low and non-significant within regions and between populations from Britain and France. For *A. digitatum*, only pairwise comparisons with North Sea populations were significant. However, for *E. verrucosa*, many pairwise comparisons between Portugal populations and populations from Britain, Ireland and France were significantly different from zero.

The PCoA suggested regional structure in *E. verrucosa* (Figure 2a), with evidence for three clusters: Portugal, Ireland, and populations from Britain and France. In contrast, little evidence of regional structure was apparent in *A. digitatum* (Figure 2b). There was some evidence for the isolation of the North Sea and UB74 Wreck populations of *A. digitatum*; however, genetic structure did not appear wholly concordant with geography, as the North Sea populations did not group together.

For the AMOVA, populations were grouped by geographical region for each species: Portugal, France, Ireland and Britain (*E. verrucosa*) and France, Ireland, Britain, and the North Sea (*A. digitatum*). In both species, global tests revealed that the majority of variation was explained by variation within populations (Supplementary Appendix 3). For *E. verrucosa*, a small but highly significant amount of variation was explained by differences between the geographical regions ($F_{CT}=0.016$, $P<0.001$). Similarly, a significant (but much smaller) amount of variation was explained by differences between regions for *A. digitatum* ($F_{CT}=0.001$, $P=0.049$).

For *E. verrucosa*, both the mean L(K) and delta K statistics indicated $K=3$ as the most probable number of discrete populations within the data set (Supplementary Appendix 4). STRUCTURE analysis (Figure 3a) identified essentially the same groupings as observed in the PCoA (Figure 2a), but also indicated that all *E. verrucosa* colonies from France (and a few from Britain) shared some allelic similarities with *E. verrucosa* from Portugal. To explore potentially finer-scale population structure (<500 km distance between sites) in populations from Britain and France, a hierarchical STRUCTURE analysis was conducted using data from only these regions. The most likely number of populations was identified as $K=2$ (Supplementary Appendix 4), which revealed moderate structure partitioned between *E. verrucosa* populations from Britain and those from France, with some evidence of allelic variants more typical of *E. verrucosa* from France occurring in samples from Britain (Figure 3b). In contrast, for *A. digitatum*, the mean L(K) suggested panmixia ($K=1$; Supplementary Appendix 4). Analysis of delta K for *A. digitatum* suggested $K=2$; however, the delta K method is known to be unsuitable for accurately identifying K when $K=1$ (Evanno *et al.*, 2005).

Analysis of pairwise genetic and geographic distances between sample sites showed a moderate, significant correlation for *E. verrucosa* ($r^2=0.348$, $P=0.001$; Figure 4a). The correlation was much weaker but remained significant when the samples from Portugal were excluded from the analysis ($r^2=0.083$, $P=0.004$; Supplementary Appendix 5). Similarly, the correlation remained significant when the samples from Portugal and Ireland were excluded from the analysis ($r^2=0.077$, $P=0.003$) (Supplementary Appendix 5). For *A. digitatum*, a weak, but similarly significant correlation between

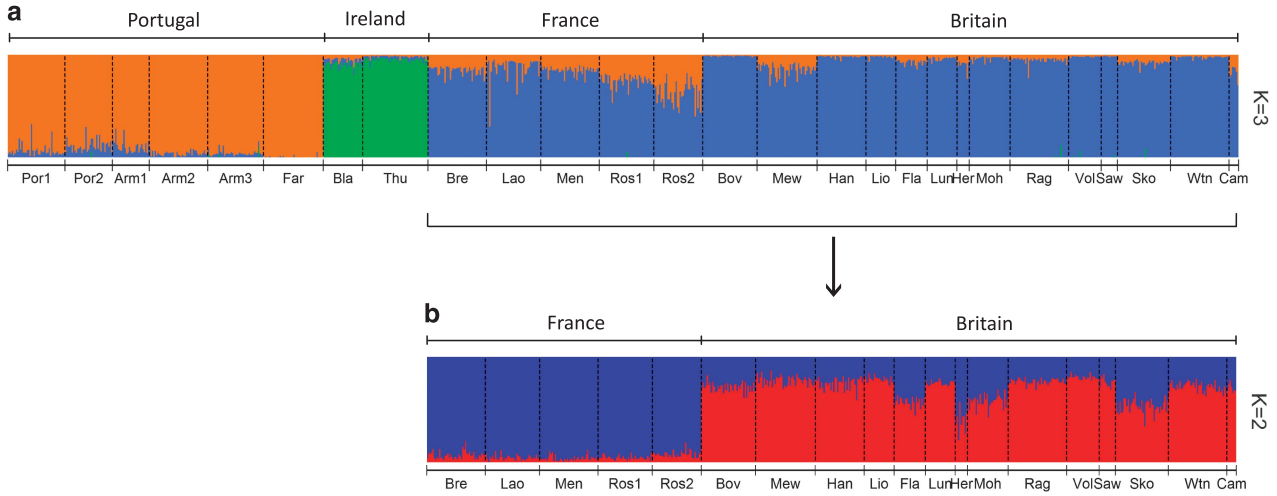


Figure 3 STRUCTURE analysis for *Eunicella verrucosa* using all populations (**a**) and hierarchical STRUCTURE analysis using populations from only Britain and France (**b**). The colours in the STRUCTURE plots correspond to genetic clusters, in which each individual is represented as a coloured vertical bar that represents that individual's membership in each cluster. See Table 1 for details on population codes.

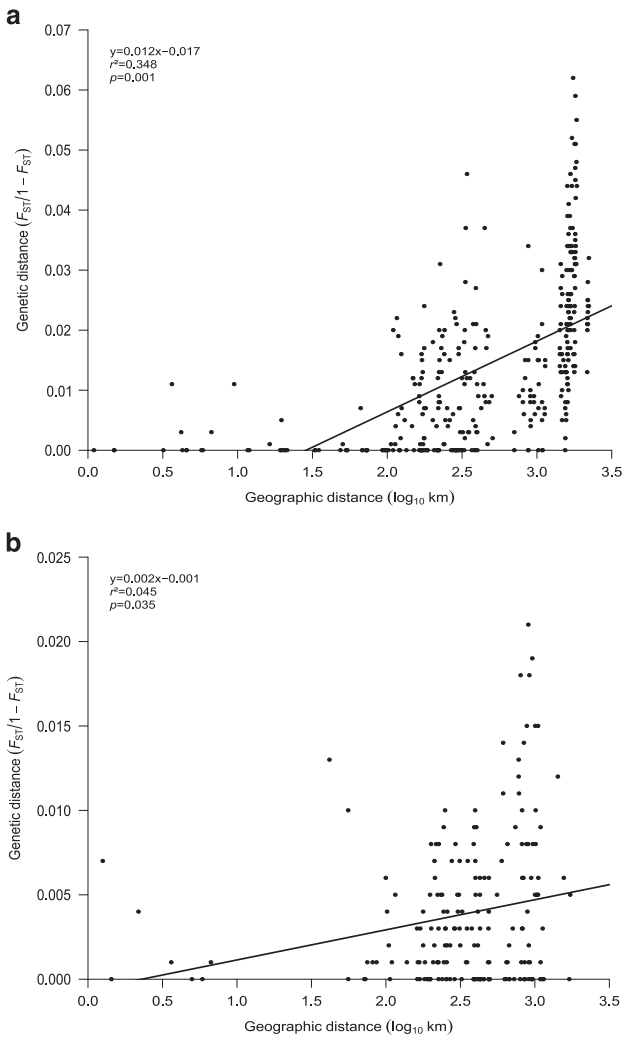


Figure 4 Relationship between genetic distance and geographic distance for *Eunicella verrucosa* (**a**) and *Alcyonium digitatum* (**b**).

genetic and geographic distances was apparent ($r^2 = 0.045$, $P = 0.035$; Figure 4b); however, removal of the North Sea samples resulted in no correlation ($r^2 < 0.001$, $P = 0.463$; Supplementary Appendix 5). Analysis of both species was also carried out using G''_{ST} as the genetic distance; for *E. verrucosa* the result was similar to that obtained using F_{ST} (Supplementary Appendix 5), however, for *A. digitatum*, the correlation was lower and non-significant (Supplementary Appendix 5).

Gene flow and effective population size

To estimate gene flow, samples of both species were classified by geographical region as per the AMOVA groupings. Contemporary gene flow estimates (using BayesAss) for both species indicated that the majority of gene flow originated from sample sites within regions (Figure 5). However, for both species, where some gene flow between regions was detected, populations in southwest Britain were the predominant source of exogenous allelic variants. For *E. verrucosa*, gene flow from Britain was predominantly into France, whereas in *A. digitatum* gene flow from southwest Britain into the North Sea, Ireland and France was observed. In comparison, contemporary gene flow into Britain appeared very limited for both species. For *E. verrucosa*, little genetic material was exchanged between Ireland and any other region; likewise, gene flow to/from Portugal was minimal, except for some minor gene flow from Portugal into France. Little or no gene flow from France was detected, suggesting that *E. verrucosa* in both France and Ireland are effectively sinks. In contrast, for *A. digitatum*, some gene flow from France to other regions was apparent, although gene flow from the North Sea and Ireland to other study areas was all but absent. For both species, estimates of historical gene flow (using Migrate-n) were somewhat more complex, with populations from Britain again acting as the main source of gene flow for both species, and with only limited gene flow into southwest Britain (Figure 5). Overall, historically, there appeared to have been considerably more gene flow between all regions.

Analyses of N_e were run using the same groupings as used in the gene flow analyses. Estimates of contemporary effective population sizes were infinite for both species (Supplementary Appendix 6). Historical effective population sizes for *E. verrucosa* indicated that samples from Britain had the largest N_e, followed by those from

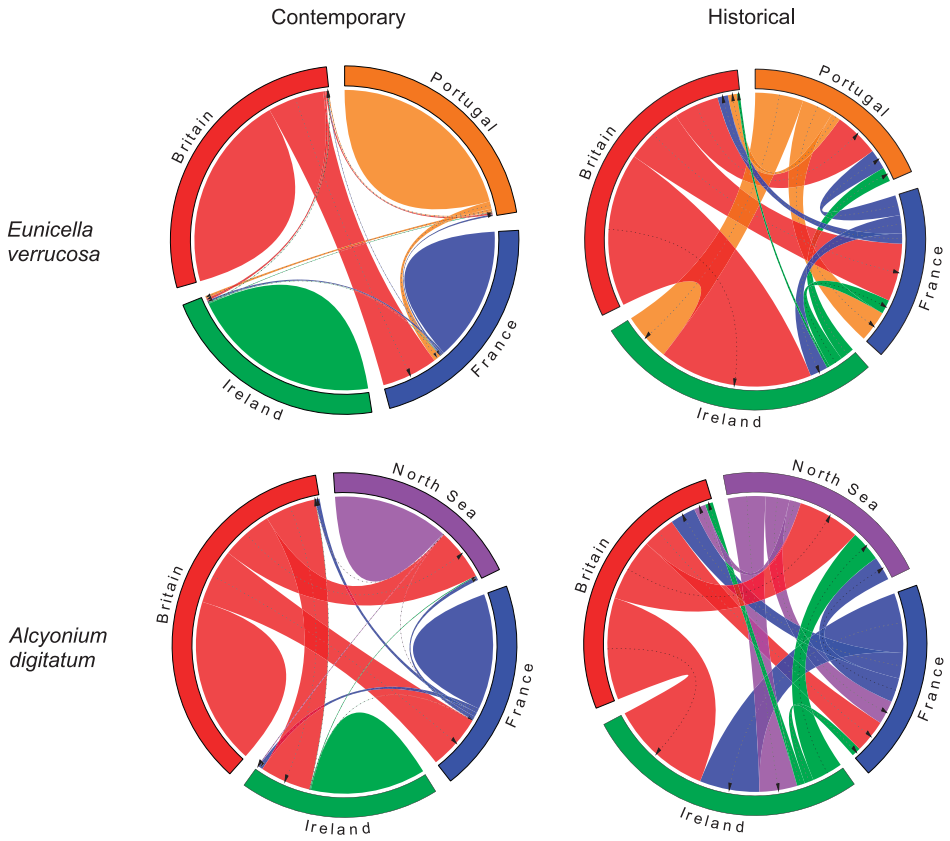


Figure 5 Gene flow diagrams for *Eunicella verrucosa* and *Alcyonium digitatum*. Contemporary gene flow estimates were derived from BayesAss and historical gene flow estimates were calculated using Migrate-n. Colours correspond to regions: Britain (red), France (blue), Ireland (green), Portugal (orange), the North Sea (purple). The direction of an arrow represents the direction of gene flow from one region to another. The width of the arrows denotes the relative amount of gene flow within the scenario being explored (that is, the wider the arrow, the more gene flow). The 'humps' in the estimates of contemporary gene flow represent gene flow originating from sample sites within regions. Patterns for each diagram are independent, that is, similar widths of arrows or humps do not represent the same amount of gene flow across each of the four diagrams (see Supporting Appendix 6 for exact gene flow estimates).

Ireland, France and Portugal. In contrast, for *A. digitatum*, estimates of historical N_e in Ireland and the North Sea were by far the largest, being more than six times larger than the N_e for *E. verrucosa* in Britain. Estimates of N_e for *A. digitatum* from Britain and France were, in contrast, very small.

DISCUSSION

This study demonstrates that regional population structure is apparent in the octocoral species *E. verrucosa* sampled from sites around the northeast Atlantic, including northwest Ireland, southwest Britain, northwest France and southern Portugal. However, over a similar spatial area, another temperate octocoral, *A. digitatum*, showed only very limited population structure. Therefore, despite the similarities in habitat and life histories of these octocorals, patterns of genetic connectivity over approximately the same geographical area appear variable between species within Octocorallia. The implications of and possible causes for these apparent differences –differences in gene flow and/or effective population size– are now considered.

Genetic diversity and inbreeding

Genetic diversity measures (H_{exp} and A_r) were generally uniform across the sampling ranges of each species (Tables 1 and 2); however, higher estimates of both measures in *A. digitatum* indicated higher genetic diversity in this species than in *E. verrucosa*. In comparison to other studies of temperate corals (Table 3), the genetic diversity of *A. digitatum* observed in the current study was higher than or comparable to that reported in the octocorals *Eunicella singularis* (Costantini *et al.*, 2016) and *E. cavolini* (Masmoudi *et al.*, 2016), and the stony coral *Astroides calycularis* (Casado-Amezú *et al.*, 2012), but less than two other Mediterranean octocorals, *Corallium rubrum* (Ledoux *et al.*, 2010) and *Paramuricea clavata* (Mokhtar-Jamai *et al.*, 2011). In contrast, *E. verrucosa* exhibited the lowest genetic diversity, a finding that may be explained by both biological/ecological and genetic methodology factors: one highly variable locus, Ever009, which exhibited nine alleles when originally developed (Holland *et al.*, 2013a), was excluded from the current analysis due to the presence of null alleles. At the same time, while the low diversity statistics reflect low polymorphism at some *E. verrucosa* loci, reduced polymorphism may itself have been the product of an overall lower level of genetic diversity within the populations studied: at four loci (Ever005, Ever008, Ever011 Ever012) only two or three alleles were detected during initial testing (Holland *et al.*, 2013a), with a maximum of five alleles detected at these loci in the current study. The precise biological/ecological causes of this low genetic diversity (for example, inbreeding, selection) remain to be determined. Overall, differences in the patterns of genetic diversity (H_{exp} and A_r) detected between the two species studied were markedly consistent and may, at least in part, be explained by higher genetic connectivity in *A. digitatum*.

Eunicella verrucosa has previously been reported as having a low dispersal potential (Munro, 2004); if correct, this would increase

Table 3 Summary table comparing the two octocoral species in the current study with previous studies of temperate corals that used microsatellite markers to investigate genetic diversity and population structure in the northeast Atlantic (Atl) and the Mediterranean (Med)

Species information				*Study information*						
Species	*Reproduction*	*PLD*	*Clonality*	*Sea S; N*	*No. loci*	*Mean H_{exp}*	*Mean A_r*	*Structure*	*IBD*	*References*
Soft coral										
Alcyonium digitatum	Broadcast	Lecithotrophic, unknown	Yes	Atl 20; 648	8	0.63	4.18	No, $F_{ST}=0.003$	Weak, $r^2=0.05$	This study; Hartnoll, 1975
Eunicella verrucosa	Broadcast	Lecithotrophic, unknown	Limited	Atl 27; 905	13	0.42	2.58	Yes, $F_{ST}=0.012$	Yes, $r^2=0.35$	This study; Munro, 2004
Eunicella singularis	Brooder	Unknown	Unknown	Med 13; 301	6	0.53	3.58	Yes, NA	NA	Costantini *et al.*, 2016
Eunicella cavolini	Unknown	Unknown	Unknown	Med 19; 584	7	0.56	4.24	Yes, $F_{ST}=0.130$	Yes, $r^2=0.57$	Masmoudi *et al.*, 2016
Corallium rubrum	Brooder	Lecithotrophic, 4–12 days	Limited	Med 40; 1222	10	0.74	7.30	Yes, $F_{ST}=0.097$	Yes, $r^2=0.15$	Ledoux *et al.*, 2010
Paramuricea clavata	Surface brooder	Lecithotrophic, 6–23 days	Unknown	Med 39; 1114	6	0.74	6.48	Yes, $F_{ST}=0.116$	Yes, $r^2=0.51$	Mokhtar-Jamai *et al.*, 2011
Stony coral										
Astroides calycularis	Brooder	Unknown	Unknown	Med 16; 381	13	0.55	4.38	Yes, $F_{ST}=0.236$	Yes, $r^2=0.45$	Casado-Amezúa *et al.*, 2012

Abbreviations: A_r, allelic richness; H_{exp}, expected heterozygosity; *IBD*, isolation by distance; *N*, number of individuals genotyped; NA, information not available; *PLD*, pelagic larval duration; *S*, number of sites.

the potential for inbreeding. However, the findings of Munro (2004) were based on analysis of only four isoenzymes, markers notorious for their lack of resolution compared to more modern PCR-based techniques (for example, Stevens and Tibayrenc, 1995) and the range and limited number of significant inbreeding coefficients (F_{IS}) observed for *E. verrucosa* in the current study suggests that the frequency of inbreeding is low, variable between sites and likely due to site-specific factors. Regarding the use of F_{IS}, while the coefficient is typically referred to as measuring the degree of inbreeding within a population, it actually measures homozygosity excess relative to Hardy-Weinberg expectations, and other processes, for example, the inadvertent combining of data from populations with different allele frequencies (the so called 'Wahlund effect') can also drive significant positive F_{IS} results. Such a consideration is relevant when seeking to explain the higher number of significant positive F_{IS} values obtained for *E. verrucosa* populations (Table 1), as this species showed considerably more evidence of genetic structuring (Figure 2a, Supplementary Appendix S2b) than did *A. digitatum* (Figure 2b, Supplementary Appendix S2d) across the range studied. Thus, given the higher proportion of significant between-population pairwise F_{ST}s (Supplementary Appendix S2b) observed for *E. verrucosa*, it is possible that cryptic intra-population genetic differentiation may also have played a role in driving significant F_{IS} values in this species. If our F_{IS} results (especially for *E. verrucosa*) were due to Wahlund effects, such findings would suggest even less inbreeding within the species than the small amount currently postulated. Additionally, the generally low F_{IS} values observed also accord with the low proportion of duplicate genotypes detected in *E. verrucosa* ($<2\%$) across the study; the number of *E. verrucosa* individuals at a site with duplicate genotypes ranged from 0 (most samples) to 4 in a small sample (Cam, $N=11$) from north Cornwall. Interestingly, the north Cornwall sample was one of the few samples not collected by our dive teams, and the relatively high proportion of duplicate genotypes at this site may be a reflection of sampling practice rather than biological reality.

For *A. digitatum*, significant F_{IS} coefficients were even fewer and lower (though still positive), suggesting only very limited inbreeding in this species; likewise, only a very low proportion (~1%) of all individual *A. digitatum* successfully genotyped had duplicate profiles. While such findings might be expected for broadcast spawning corals (Ayre and Hughes, 2000), exceptions to this pattern are not uncommon; for example, Combosch and Vollmer (2011) studied populations of *Pocillopora damicornis*, a broadcast spawning tropical reef coral, and reported a range of large, mostly positive, significant inbreeding coefficients (F_{IS} range: -0.048–0.421), leading them to conclude that widespread inbreeding was apparent in this species in the eastern Pacific.

Compared to previous population genetics studies in octocoral species (for example, Ledoux *et al.*, 2010; Mokhtar-Jamai *et al.*, 2011), the number of significant F_{IS} estimates reported here for *E. verrucosa* and *A. digitatum* is globally low: ten significant F_{IS} estimates at 27 sites for *E. verrucosa* (Table 1) and four significant F_{IS} estimates at 20 sites for *A. digitatum* (Table 2). Overall, such low estimates of F_{IS}, considered together with the very low numbers of identical individuals sampled in both species, is suggestive of low levels of inbreeding in these two species of octocoral in these parts of their respective ranges.

Genetic structure and connectivity

Population genetic structure was apparent for *E. verrucosa* at a regional spatial scale (500–2000 km between sample sites), suggesting

restrictions to gene flow between populations in different geographical regions. This finding was also supported by analyses of contemporary gene flow (though less so historically), as demonstrated by the limited exchange of genetic material between regions, except for some gene flow between Britain and France (<500 km distance between sites; Figures 3b and 5). Indeed, in both species, analysis of contemporary gene flow suggested that the majority of gene flow occurred between sites within geographical regions, as also observed in *Paramuricea clavata*, a Mediterranean octocoral (Arizmendi-Mejía *et al.*, 2015). In contrast to *E. verrucosa*, little regional structure was apparent in *A. digitatum*, and the only significant differentiation detected was between the samples from the North Sea and those from more westerly areas (>550 km distance between sites); this differentiation appeared to be the product of IBD (as evidenced when comparing the results of IBD analysis with and without North Sea samples of *A. digitatum*; see Figure 4b and Supplementary Figure S5b) and/or a barrier between the samples of western origin and those from the North Sea. Estimates of contemporary effective population size (N_e) were infinite for both species (Supplementary Appendix 6). Assuming these estimates to be accurate, we found no evidence for disequilibrium caused by genetic drift due to a finite number of parents and, thus, any disequilibrium observed was due to sampling error (Waples and Do, 2008). In contrast, estimates of historical effective population sizes were smaller and variable between regions (Supplementary Appendix 6); this result, together with findings from the corresponding analyses of historical and contemporary gene flow (Figure 5) suggest historical patterns of connectivity were not the same as those observed today.

Overall, our findings suggest that *A. digitatum* is panmictic across the western part of the sampled range. One possible explanation for this apparent panmixia is that the winter spawning of *A. digitatum* may facilitate longer dispersal distances via wind-driven currents, thereby increasing genetic connectivity in the eastern Atlantic. Panmixia across similar spatial scales has been reported previously in other marine taxa, including cuttlefish (Wolfram *et al.*, 2006; microsatellite-based study), sea stars (Baus *et al.*, 2005; AFLP-based study), and a closely related species, *Alcyonium hibernicum* (McFadden, 1999; isoenzyme-based study), although in the latter study, in which little or no genetic variation was detected in *A. hibernicum* across the Atlantic, McFadden (1999) also linked her findings to high levels of asexual reproduction by parthenogenesis. Similarly, a recent broad study by Gagnaire *et al.* (2015) highlights the potential impact of large effective population size as an alternative explanation to contemporary panmixia in acting to limit genetic drift, thereby constraining the development of genetic structure, even where gene flow is restricted.

For *E. verrucosa*, populations from Portugal were differentiated from the majority of populations north of the Bay of Biscay. This may represent a natural break in gene flow in which genetic drift either side of the break is the primary driver of population structure; such a conclusion is supported by both the multivariate (PCoA) and Bayesian clustering (STRUCTURE) analyses. This pattern has been reported previously in a number of taxa, including bivalves (Arias *et al.*, 2010), brittlestars (Muths *et al.*, 2009), crustaceans (Papetti *et al.*, 2005; Remerie *et al.*, 2009), microturbellarians (Casu *et al.*, 2011), macroalgae (Neiva *et al.*, 2014) and fish (Milano *et al.*, 2014). However, the significant correlation between genetic and geographic distances in *E. verrucosa* in the current study indicates that a proportion of the differentiation observed is likely explained by IBD. Further analysis omitting the Portugal populations suggested that IBD explains some of the genetic differentiation observed between Portugal and all other populations, but much less of the differentiation observed between Britain, Ireland and France. Interestingly, comparisons with other temperate corals (Table 3) suggest that contemporary patterns of population structure appear often to be driven, at least in part, by IBD, which is possibly due to their sedentary life history and their lack of or shorter pelagic larval duration compared to other benthic marine species. In *E. verrucosa*, analysis of IBD showed no change in significance when the samples from Ireland were removed (Supplementary Appendix 5 and Supplementary Figure S5b), indicating IBD to be a less important driver of population structuring in these Irish samples; such a finding suggests genetic differentiation of these range-peripheral populations is more likely driven by other factors, for example, barriers to gene flow and genetic drift and/or selection. Several previous studies of invertebrates sampled from across this region have also reported genetic differentiation in western Ireland compared to other locations in the northeast Atlantic (Remerie *et al.*, 2009; Casu *et al.*, 2011). These studies attributed this differentiation to recolonisation from relatively northerly refugia that persisted in ice-free coastal areas during the last glacial maximum; however, while Casu *et al.* explained their findings (reduced genetic diversity in more northerly recolonized populations) by reference to founder effects and low numbers of recolonisers (Hewitt, 1996,1999), Remerie *et al.* postulated the higher genetic diversity and heterogeneity they observed in glaciated areas to be suggestive of range persistence during the last glacial maximum. Our findings for *E. verrucosa* from Ireland (which exhibited the lowest genetic diversity detected in our entire study [H_{exp}, A_r]) are in line with those of Casu *et al.* (2011) and, likewise, are suggestive of founder effects following post-glacial recolonisation of suitable northerly habitats by small numbers of recolonisers (Nichols and Hewitt, 1994). A lack of sampling at the southern-most limits of the range of *E. verrucosa* also makes it difficult to infer the precise origins of the populations in northwest Ireland, as gaps in our knowledge concerning the genetic identity of all possible source populations limits the accuracy of any putative recolonisation hypotheses. Furthermore, to what degree the contemporary distribution of *E. verrucosa* reflects the extent of the species at the last glacial maximum is unknown, but, to date, its distribution appears not to have extended to areas known to be under ice during the last glacial maximum (Hayward and Ryland, 1995; Hewitt, 1996). In contrast, the distribution of *A. digitatum* in the northeast Atlantic does not show the same pattern and its present day distribution is considerably more northerly, extending from northern Iberia and the Bay of Biscay up to Iceland and Norway (Hayward and Ryland, 1995). Another possible explanation for the differentiation observed in *E. verrucosa* from Ireland in the current study is that the effect of selection may be sufficiently strong in northwest Ireland to mitigate the homogenising effect of gene flow. The populations of *E. verrucosa* found in northwest Ireland are known to be peripheral and inhabit the most northerly limits of the species range (Hayward and Ryland, 1995). Moreover, the lower measures of expected heterozygosity and allelic richness observed in both Irish samples are characteristic of marginal populations, which typically have reduced genetic diversity and can often be under intense selection pressures (Johannesson and André, 2006); our tests for selection identified at least one locus under positive selection. At this stage, however, we do not know which selection pressures, if any, may be acting on these most northerly populations of pink sea fan.

In contrast to the patterns observed between regions, our findings for both octocoral species suggested high gene flow and/or large effective population sizes within regions. For example, for *E. verrucosa*, little differentiation was observed between the two most distant populations within Britain (Sawtooth and Skomer), implying that the transfer of genetic material can potentially occur up to distances of ~480 km. For *A. digitatum*, gene flow was evident at an even larger spatial scale, suggesting that genetic material can be transferred greater distances, potentially more than 1050 km (Payne's Rock—Norfolk). These results suggest that genetic connectivity is high at an intra-regional scale in both species. However, as observed in many marine species with similar life history traits, large effective population size can also act to reduce (or eliminate) genetic structure, sometimes even in situations were gene flow is limited (Gagnaire *et al.*, 2015). Thus, in postulating high gene flow within regions, we need to be mindful of the potential effects of large effective population sizes on genetic structure (or lack of) in these two species.

The hierarchical analysis of *E. verrucosa* population structure revealed a small degree of genetic differentiation between populations in southwest Britain and northwest France at a distance (~200 km) less than that separating some British populations; however, minimal differentiation was evident for *A. digitatum* across this area. The effects of mid-channel currents and local near-shore eddies (Dauvin, 2012) on cross-channel larval migration remains to be explored, although previous research has identified a potential genetic break around western Brittany in a number of taxa, including polychaetes (Jolly *et al.*, 2005), nematodes (Wielgoss *et al.*, 2008) and bivalves (Becquet *et al.*, 2012). In this study, the contrast in genetic connectivity across the English Channel may result from differences in the reproductive biology of the two study species. The pelagic larval duration of *E. verrucosa* is not known, however, evidence from this study suggests this could be shorter than the pelagic larval duration for *A. digitatum*.

Conservation implications and MPAs

Eunicella verrucosa has been listed under the IUCN red list Vulnerable A1d category since 1996, and is recognised as a species facing a high- to medium-term extinction risk due to exploitation (International Union for Conservation of Nature and Natural Resources (IUCN), 2017). It is also listed as a priority species under the UK Biodiversity Action Plan, the UK response to the prevention of biodiversity loss called for by the 1992 Convention on Biological Diversity; in the Republic of Ireland, France and Portugal it does not currently receive any additional protection beyond its IUCN listing. Several of the MCZs recently designated around southwest Britain (for example, Chesil Beach and Stennis Ledges, The Manacles, and The Isles of Scilly) specifically identify *E. verrucosa* as a Protected Feature in their designation listing, and 60% of *E. verrucosa* colonies recorded by diver surveys in southwest Britain fall within areas protected by various other European Union legislation (Pikesley *et al.*, 2016). However, not all of these areas are protected from bottom trawling (for example, The Manacles, Whitsand Bay, Chesil Beach and Stennis Ledges MCZs), suggesting that a large proportion of *E. verrucosa* in Britain remains vulnerable to anthropogenic disturbance and the current level of protection of UK marine ecosystems afforded by the MCZ network is generally insufficient (for example, Lieberknecht and Jones, 2016; Pikesley *et al.*, 2016). Moreover, while the UK government appears to have moved away from the recommended ecological network guidelines for MCZ designation (Lieberknecht and Jones, 2016), the *E. verrucosa* data presented here highlight interesting findings relative to the conservation of ecologically important and prevalent sessile taxa at local (that is, single-site MPAs) to regional (that is, connected metapopulation) scales.

More specifically, the genetic distinctiveness of *E. verrucosa* populations from Ireland underpins an argument for protecting particular sites. Marginal populations often contain rare alleles (the highest extent of private alleles were found at these sites), but may recruit more slowly, and may be genetically isolated, implying vulnerability and reduced resilience (Sanderson, 1996) and therefore an increased need for protection. However, away from the edges of the species range, our data suggest that connectivity can be maintained between populations of these species in some designated MCZs. Moreover, the range of *E. verrucosa* in Britain is small compared to its (primarily Lusitanean) global distribution and, although contemporary connectivity between British populations appears to be a high, at regional spatial scales it could be argued that the genetic distinction of these populations, coupled with their possible role as source populations that act to maintain broader connectivity across this area of the northeast Atlantic, may be sufficient to warrant international conservation efforts (for example, OSPAR Convention, 2013).

In the UK, *A. digitatum* has no specific protective status, is not at the periphery of its global range, and, in our study (apart from the North Sea—English Channel/eastern Atlantic divide), it exhibited relatively high genetic diversity with little evidence of any major barriers to gene flow. Overall, coupled with the high prevalence of this species in UK waters, these factors imply that this species may be a low priority for protection in its own right, and it is likely to receive only patchy, incidental protection based upon the location of current MCZs designated on the basis of other features, although arguments for consideration of this species in design guidelines could still fit both the 'representativity' and 'replication' principles (Natural England, 2010). However, reduced heterozygosity and impaired sexual reproduction have been reported in another cnidarian species subjected to trawling damage (Henry and Kenchington, 2004) and reduced colony numbers and size have been reported for *A. digitatum* in Lyme Bay, southern England, in trawled areas (Hinz *et al.*, 2011); therefore, this species may be locally vulnerable. Certainly, the occurrence of damaged, sessile populations in disturbed areas, are a useful proxy to highlight degraded ecosystems that may also contain more directly threatened species.

The results from the present study suggest that populations of *E. verrucosa* would benefit from protection across the species range as a connected metapopulation. Although implementing protective measures for a single species across its entire UK range is highly unlikely given commercial and economic pressures within the region, our study serves to highlight areas for consideration in an ecosystem-based management approach. Irish populations of *E. verrucosa* may warrant protection because of their marginality, yet they are not currently protected within the Republic of Ireland beyond their IUCN listing. Analysis of gene flow for both octocoral species studied suggests populations in southwest Britain act as a source for surrounding regions, highlighting the value in protecting these populations. In the UK, the current recommendation for the spacing of designated MPAs is in the region of 40–80 km (Natural England, 2010). In light of our findings, it appears that the distances between these MPAs would generally be sufficient to maintain genetic connectivity of these two octocoral species in UK waters. Of course, this assumes that contemporary local oceanic currents are able to facilitate the transport of enough larvae in each species, whether by a continuing stepping-stone process or a single dispersal event. Managing effective conservation of marine species with overlapping generations and high levels of clonality,

such as sponges and corals, can be challenging because characteristic genotypes may persist for decades to centuries, even after significant barriers to gene flow arise. As a result, traditional *F*-statistics may not always represent current patterns of genetic connectivity (Botsford *et al.*, 2009) and these factors should be incorporated when including genetic data into MPA network designation. Furthermore, because of the challenges associated with genotyping octocorals, such as the slow rate of mitochondrial evolution (McFadden *et al.*, 2010) and the difficulty of isolating microsatellites (Liu *et al.*, 2005), the type and numbers of molecular marker used may not be powerful enough to detect a signal of fine-scale population structure. As seen in the current study, while relatively strong patterns of regional structure were detected in *E. verrucosa*, except for some weak structuring between French and English samples of *E. verrucosa* in the Channel, no fine-scale structure (<200 km between sample sites) was detected for either of the species studied. This may be indicative of genetic connectivity between these populations, but could also represent a lack of power in the genetic markers used. Exploration of alternative marker systems may deliver improved resolution (for example, Shinzato *et al.*, 2015) and should prove valuable for future conservation research and the management of MPA networks.

In conclusion, genetic diversity appears to be uniform across the range studied in both species; however, genetic diversity was low in *E. verrucosa*, whereas in *A. digitatum*, it was slightly higher, but still lower than that reported for two species of Mediterranean octocorals (Ledoux *et al.*, 2010; Mokhtar-Jamai *et al.*, 2011). For both species, only limited inbreeding was apparent, and whether this has an impact on fitness and long-term resilience of the populations in question is currently unknown. Regional population structure was identified in *E. verrucosa*, indicative of departures from panmixia at large spatial scales; in contrast, in *A. digitatum*, apart from some genetic differentiation between populations from the North Sea and those from the English Channel/eastern Atlantic, we found little population structure, suggesting high gene flow and connectivity in this species in the western part of the range sampled. Contemporary and historical estimates of effective population size were contrasting and generally difficult to interpret, and for both species the potential role of large N_es in masking a lack of gene flow cannot be ruled out. Patterns of gene flow were complex, but indicated Britain as a source of genetic variants for both species. Several populations of both species are represented in the UK MPA network and, given the ecological importance of both species, continued monitoring and assessment of their genetic diversity within and beyond protected sites could be a useful measure of the efficacy of the existing network, and a valuable guide to the designation of new MCZs.

DATA ARCHIVING

Microsatellite genotypes for *Eunicella verrucosa* and *Alcyonium digitatum* are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.nj0v4.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank D Laffoley (Natural England, IUCN) and J Ashworth (Natural England) for invaluable advice and help with project conception and planning. We are grateful to the following for assistance with sample collection: J Kent and S Williams (University of Bristol Underwater Club), P Shears and J Shears (University of Exeter), J Turner (University of Exeter Sub-Aqua Club), J Hall-Spencer and M Whelan (Plymouth University), K Hiscock (Marine Biological Association UK), P Newman (Countryside Council for Wales), C Munro (Marine Bio-images, Exeter), R Jones (Zoological Society of London), I Reach (Natural England), B Pearce and J Pinnion (Marine Ecological Surveys), P Romans and staff (Observatoire Océanologique de Banyuls, France), L Lévêque and staff (Station Biologique de Roscoff, France) and R Castilho and staff (University of Algarve, Portugal). Sampling within the UK was carried out under Natural England licences 20080861 and 20090943, and Marine Management Organisation licence MMO-0001. Genotyping was performed at the Natural Environment Research Council, UK (NERC) Biomolecular Analysis Facility, University of Sheffield and we thank D Dawson, G Horsburgh and A Krupa for technical assistance. We thank Natural England (project No. RP0286, contract No. SAE 03-02-146), the NERC (grant No. NE/L002434/1) and the University of Exeter for funding this research. Additional funding for sample collection, travel and microsatellite development was provided by the EU Framework 7 ASSEMBLE programme, agreement no. 227799, and NERC grant No. NBAF-362. We thank the editor and three reviewers for their helpful comments in revising this manuscript. This research forms part of the PhDs of LPH and TLJ. LPH is currently at the School of Biological Sciences, Victoria University of Wellington, New Zealand.

AUTHOR CONTRIBUTIONS

LPH conducted laboratory analysis, carried out sample collection, undertook preliminary analyses and wrote the manuscript. TLJ analysed the data and wrote the manuscript. JRS conceived, designed and directed the study, carried out sample collection and wrote the manuscript.

Almany GR, Connolly SR, Heath DD, Hogan JD, Jones GP, McCook LJ *et al.* (2009). Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs. *Coral Reefs* **28**: 339–351.

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008). LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics* **9**: 323.

Arias A, Fernández-Moreno M, Fernández-Tajes J, Gaspar MB, Méndez J (2010). Strong genetic differentiation among east Atlantic populations of the sword razor shell (*Ensis siliqua*) assessed with mtDNA and RAPD markers. *Helgoland Mar Res* **65**: 81–89.

Arizmendi-Mejía R, Linares C, Garrabou J, Antunes A, Ballesteros E, Cebrian E, Díaz D, Ledoux J-B (2015). Combining genetic and demographic data for the conservation of a mediterranean marine habitat-forming species. *PLoS One* **10**: e0119585.

Aurelle D, Ledoux J-B, Rocher C, Borsa P, Chenuil A, Féral J-P (2011). Phylogeography of the red coral (*Corallium rubrum*): inferences on the evolutionary history of a temperate gorgonian. *Genetica* **139**: 855–869.

Ayre DJ, Hughes TP (2000). Genotypic diversity and gene flow in brooding and spawning corals along the Great Barrier Reef, Australia. *Evolution* **54**: 1590–1605.

Bargelloni L, Alarcon JA, Alvarez MC, Penzo E, Magoulas A, Palma J *et al.* (2005). The Atlantic-Mediterranean transition: discordant genetic patterns in two seabream species, *Diplodus puntazzo* and *Diplodus sargus*. *Mol Phylogenet Evol* **36**: 523–535.

Bargelloni L, Alarcon JA, Alvarez MC, Penzo E, Magoulas A, Reis C *et al.* (2003). Discord in the family *Sparidae* (*Teleostei*): divergent phylogeographical patterns across the Atlantic-Mediterranean divide. *J Evol Biol* **16**: 1149–1158.

Baus E, Darrock DJ, Bruford MW (2005). Gene-flow patterns in Atlantic and Mediterranean populations of the Lusitanian sea star *Asterina gibbosa*. *Mol Ecol* **14**: 3373–3382.

Becquet V, Simon-Bouhet B, Pante E, Hummel H, Garcia P (2012). Glacial refugium versus range limit: conservation genetics of *Macoma balthica*, a key species in the Bay of Biscay (France). *J Exp Mar Bio Ecol* **432–433**: 73–82.

Beerli P, Felsenstein J (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc Natl Acad Sci USA* **98**: 4563–4568.

Botsford LW, White JW, Coffroth MA, Paris CB, Planes S, Shearer TL *et al.* (2009). Connectivity and resilience of coral reef metapopulations in marine protected areas: matching empirical efforts to predictive needs. *Coral Reefs* **28**: 327–337.

Casado-Amezú P, Goffredo S, Templado J, MacHordom A (2012). Genetic assessment of population structure and connectivity in the threatened mediterranean coral *Astroides calycularis* (*Scleractinia, Dendrophylliidae*) at different spatial scales. *Mol Ecol* **21**: 3671–3685.

Casu M, Sanna D, Cossu P, Lai T, Francalacci P, Curini-Galletti M (2011). Molecular phylogeography of the microturbellarian *Monocelis lineata* (Platyhelminthes: Proseriata) in the North-East Atlantic. *Biol J Linn Soc* **103**: 117–135.

Charrier G, Chenel T, Durand JD, Girard M, Quiniou L, Laroche J (2006). Discrepancies in phylogeographical patterns of two European anglerfishes (*Lophius budegassa* and *Lophius piscatorius*). *Mol Phylogenet Evol* **38**: 742–754.

Combosch DJ, Vollmer SV (2011). Population genetics of an ecosystem-defining reef coral *Pocillopora damicornis* in the Tropical Eastern Pacific. *PLoS One* **6**: e21200.

Costantini F, Fauvelot C, Abbiati M (2007). Genetic structuring of the temperate gorgonian coral (*Corallium rubrum*) across the western Mediterranean Sea revealed by microsatellites and nuclear sequences. *Mol Ecol* **16**: 5168–5182.

Costantini F, Gori A, Lopez-González P, Bramanti L, Rossi S, Gili J-M *et al*. (2016). Limited genetic connectivity between gorgonian morphotypes along a depth gradient. *PLoS One* **11**: e0160678.

Cowen RK, Gawarkiewicz G, Pineda J, Thorrold SR, Werner FE (2007). Population connectivity in marine systems. *Oceanography* **20**: 14–21.

Cowen RK, Sponaugle S (2009). Larval dispersal and marine population connectivity. *Ann Rev Mar Sci* **1**: 443–466.

Da Silva AG, Appleyard SA, Upston J (2014). Establishing the evolutionary compatibility of potential sources of colonizers for overfished stocks: a population genomics approach. *Mol Ecol* **24**: 564–579.

Dauvin JC (2012). Are the eastern and western basins of the English Channel two separate ecosystems? *Mar Poll Bull* **64**: 463–471.

Dunn OJ (1961). Multiple comparisons among means. *J Am Stat Assoc* **56**: 52–64.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**: 2611–2620.

Excoffier L, Lischer HEL (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* **10**: 564–567.

Fogarty MJ, Botsford LW (2007). Population connectivity and spatial management of marine fisheries. *Oceanography* **20**: 112–123.

Foster NL, Paris CB, Kool JT, Baums IB, Stevens JR, Sanchez JA *et al*. (2012). Connectivity of Caribbean coral populations: complementary insights from empirical and modelled gene flow. *Mol Ecol* **21**: 1143–1157.

Francis RM (2017). POPHELPER: an R package and web app to analyse and visualise population structure. *Mol Ecol Resour* **17**: 27–32.

Gagnaire P-A, Broquet T, Aurelle D, Viard F, Souissi A, Bonhomme F, Arnaud-Haond S, Bierne N (2015). Assessing connectivity of marine populations in the genomic era. *Evol Appl* **8**: 769–786.

Gerber LR, Del Mark Mancha-Cisneros M, O'Connor M, Selig ER (2014). Climate change impacts on connectivity in the ocean: implications for conservation. *Ecosphere* **5**: 1–18.

Goffredo S, Di Ceglie S, Zaccanti F (2009). Genetic differentiation of the temperate-subtropical stony coral *Leptopsammia pruvoti* in the Mediterranean Sea. *Isr J Ecol Evol* **55**: 99–115.

Hall-Spencer JM, Pike J, Munn CB (2007). Diseases affect cold-water corals too: *Eunicella verrucosa* (Cnidaria: Gorgonacea) necrosis in SW England. *Diseases of Aquatic Organisms* **76**: 87–97.

Hartnoll RG (1975). The annual cycle of *Alcyonium digitatum*. *Estuar Coast Mar Sci* **3**: 71–78.

Hayward PJ, Ryland JS (1995). *Handbook of the Marine Fauna of North-west Europe*. Oxford University Press: Oxford.

Hedgecock D, Barber PH, Edmands S (2007). Genetic approaches to measuring connectivity. *Oceanography* **20**: 70–79.

Henry LA, Kenchington ELR (2004). Ecological and genetic evidence for impaired sexual reproduction and induced clonality in the hydroid *Sertularia cupressina* (*Cnidaria*: *Hydrozoa*) on commercial scallop grounds in Atlantic Canada. *Marine Biology* **145**: 1107–1118.

Herrera S, Shank TM, Sánchez JA (2012). Spatial and temporal patterns of genetic variation in the widespread antitropical deep-sea coral *Paragorgia arborea*. *Mol Ecol* **21**: 6053–6067.

Hewitt GM (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc* **58**: 247–276.

Hewitt GM (1999). Post-glacial re-colonization of European biota. *Biol J Linn Soc* **68**: 87–112.

Hinz H, Tarrant D, Ridgeway A, Kaiser MJ, Hiddink JG (2011). Effects of scallop dredging on temperate reef fauna. *Mar Ecol Prog Ser* **432**: 91–102.

Hiscock K, Sharrock S, Highfield J, Snelling D (2010). Colonization of an artificial reef in south-west England-ex-HMS 'Scylla'. *J Mar Biol Assoc UK* **90**: 69–94.

Holland LP, Dawson DA, Horsburgh GJ, Krupa AP, Stevens JR (2013a). Isolation and characterization of fourteen microsatellite loci from the endangered octocoral *Eunicella verrucosa* (Pallas, 1766). *Conserv Genet Resour* **5**: 825–829.

Holland LP, Dawson DA, Horsburgh GJ, Stevens JR (2013b). Isolation and characterization of 11 microsatellite loci from the ubiquitous temperate octocoral *Alcyonium digitatum* (Linnaeus, 1758). *Conserv Genet Resour* **5**: 767–770.

International Union for Conservation of Nature and Natural Resources (IUCN) (2017). Eunicella verrucosa (Broad Sea Fan, Pink Sea Fan). Available at: http://dx.doi.org/10.2305/IUCN.UK.1996.RLTS.T8262A12903486.en (accessed on 15 December 2016).

Johannesson K, André C (2006). Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. *Mol Ecol* **15**: 2013–2029.

Jolly MT, Jollivet D, Gentil F, Thiébaut E, Viard F (2005). Sharp genetic break between Atlantic and English Channel populations of the polychaete *Pectinaria koreni*, along the north coast of France. *Heredity* **94**: 23–32.

Jones GP, Srinivasan M, Almany GR (2007). Population connectivity and conservation of marine biodiversity. *Oceanography* **20**: 100–111.

Kalinowski ST (2005). HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* **5**: 187–189.

Kalinowski ST, Taper ML, Marshall TC (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* **16**: 1099–1106.

Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013). diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol Evol* **4**: 782–788.

Kool JT, Moilanen A, Treml EA (2013). Population connectivity: recent advances and new perspectives. *Landsc Ecol* **28**: 165–185.

Ledoux JB, Mokhtar-Jamai K, Roby C, Feral JP, Garrabou J, Aurelle D (2010). Genetic survey of shallow populations of the Mediterranean red coral [*Corallium rubrum* (Linnaeus, 1758)]: new insights into evolutionary processes shaping nuclear diversity and implications for conservation. *Mol Ecol* **19**: 675–690.

Lieberknecht LM, Jones PJS (2016). From stormy seas to the doldrums: the challenges of navigating towards an ecologically coherent marine protected area network through England's Marine Conservation Zone process. *Marine Policy* **71**: 275–284.

Liu SYV, Dai CF, Fan TY, Yu HT (2005). Cloning and characterization of microsatellite loci in a gorgonian coral, *Junceella juncea* (Anthozoa; *Octocorallia*; *Ellisellidae*) and its application in clonal genotyping. *Mar Biotechnol* **7**: 26–32.

Lowe WH, Allendorf FW (2010). What can genetics tell us about population connectivity? *Molecular Ecology* **19**: 3038–3051.

Marti-Puig P, Costantini F, Rugiu L, Ponti M, Abbiati M (2013). Patterns of genetic connectivity in invertebrates of temperate MPA networks. *Adv Oceanogr Limnol* **4**: 138–149.

Masmoudi MB, Chaoui L, Topçu NE, Hammami P, Kara MH, Aurelle D (2016). Contrasted levels of genetic diversity in a benthic Mediterranean octocoral: consequences of different demographic histories? *Ecol Evol* **6**: 8665–8678.

McCook LJ, Almany GR, Berumen ML, Day JC, Green AL, Jones GP *et al*. (2009). Management under uncertainty: guide-lines for incorporating connectivity into the protection of coral reefs. *Coral Reefs* **28**: 353–366.

McFadden C (1999). Genetic and taxonomic relationships among northeastern Atlantic and Mediterranean populations of the soft coral *Alcyonium coralloides*. *Mar Biol* **133**: 171–184.

McFadden CS, Donahue R, Hadland BK, Weston R (2001). A molecular phylogenetic analysis of reproductive trait evolution in the soft coral genus. *Alcyonium*. *Evolution* **55**: 54–67.

McFadden CS, Hutchinson MB (2004). Molecular evidence for the hybrid origin of species in the soft coral genus *Alcyonium* (Cnidaria: Anthozoa: Octocorallia). *Molecular Ecology* **13**: 1495–1505.

McFadden CS, Sánchez JA, France SC (2010). Molecular phylogenetic insights into the evolution of Octocorallia: a review. *Integr Comp Biol* **50**: 389–410.

Meirmans PG, Hedrick PW (2011). Assessing population structure: FST and related measures. *Mol Ecol Resour* **11**: 5–18.

Milano I, Babbucci M, Cariani A, Atanassova M, Bekkevold D, Carvalho GR *et al*. (2014). Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). *Mol Ecol* **23**: 118–135.

Mokhtar-Jamai K, Pascual M, Ledoux JB, Coma R, Feral JP, Garrabou J *et al*. (2011). From global to local genetic structuring in the red gorgonian *Paramuricea clavata*: the interplay between oceanographic conditions and limited larval dispersal. *Mol Ecol* **20**: 3291–3305.

Mora C, Sale P (2011). Ongoing global biodiversity loss and the need to move beyond protected areas: a review of the technical and practical shortcomings of protected areas on land and sea. *Mar Ecol Prog Ser* **434**: 251–266.

Munday P, Leis J, Lough J, Paris C, Kingsford M, Berumen M, Lambrechts J (2009). Climate change and coral reef connectivity. *Coral Reefs* **28**: 379–395.

Munro L (2004). Determining the reproductive cycle of *Eunicella verrucosa*. *Reef Res*. RR Report 07/2004 ETR 12, p28.

Muths D, Jollivet D, Gentil F, Davoult D (2009). Large-scale genetic patchiness among NE Atlantic populations of the brittle star *Ophiothrix fragilis*. *Aquat Biol* **5**: 117–132.

Natural England (2010). Guidance on the size and spacing of Marine Protected Areas in England. Natural England Report (NECR037), ISSN 2040-5545. Natural England, Northminster House: Peterborough, UK.

Neiva J, Assis J, Fernandes F, Pearson GA, Serrão EA (2014). Species distribution models and mitochondrial DNA phylogeography suggest an extensive biogeographical shift in the high-intertidal seaweed *Pelvetia canaliculata*. *J Biogeogr* **41**: 1137–1148.

Nichols RA, Hewitt GM (1994). The genetic consequences of long distance dispersal during colonization. *Heredity* **72**: 312–317.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* **4**: 535–538.

OSPAR Commission (2006). Overview of OSPAR assessments 1998–2006. OSPAR Commission, London, UK, Publication No.: 287/2006, ISBN 1-905859-25-2 / 978-1-905859-25-2.

OSPAR Convention (2013). An assessment of the Ecological coherence of the OSPAR Network of Marine Protected Areas in 2012. OSPAR Commission, London, UK, ISBN 978-1-909159-52-5.

Papetti C, Zane L, Bortolotto E, Bucklin A, Patarnello T (2005). Genetic differentiation and local temporal stability of population structure in the euphausiid *Meganyctiphanes norvegica*. *Mar Ecol Prog Ser* **289**: 225–235.

Palumbi SR (2003). Population genetics, demographic connectivity, and the design of marine reserves. *Ecol Appl* **13**: S146–S158.

Peakall R, Smouse P (2012). GenAlEx 6.5 : Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **1**: 6–8.

Pérez-Portela R, Turon X, Bishop J (2012). Bottlenecks and loss of genetic diversity: spatio-temporal patterns of genetic structure in an ascidian recently introduced in Europe. *Mar Ecol Prog Ser* **451**: 93–105.

Pikesley SK, Godley BJ, Latham H, Richardson PB, Robson LM, Solandt J-L *et al.* (2016). Pink sea fans (*Eunicella verrucosa*) as indicators of the spatial efficacy of marine protected areas in Southwest UK coastal waters. *Mar Policy* **64**: 38–45.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Puckett BJ, Eggleston DB, Kerr PC, Luettich RA (2014). Larval dispersal and population connectivity among a network of marine reserves. *Fish Oceanogr* **23**: 342–361.

Rambaut A, Suchard M, Xie D, Drummond A (2014). Tracer v1.6. Available at: http://beast.bio.ed.ac.uk/Tracer.

R Development Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at: https://www.R-project.org/.

Remerie T, Vierstraete A, Weekers PHH, Vanfleteren JR, Vanreusel A (2009). Phylogeography of an estuarine mysid, *Neomysis integer* (Crustacea, *Mysida*), along the north-east Atlantic coasts. *J Biogeogr* **36**: 39–54.

Rousset F (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour* **8**: 103–106.

Sanderson WG (1996). Rarity of marine benthic species in Great Britain: Development and application of assessment criteria. *Aquat Conserv* **6**: 245–256.

Shinzato C, Mungpakdee S, Arakaki N, Satoh N (2015). Genome-wide SNP analysis explains coral diversity and recovery in the Ryukyu Archipelago. *Sci Rep* **5**: 18211.

Storey JD, Tibshirani R (2003). Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* **100**: 9440–9445.

Stevens JR, Tibayrenc M (1995). Detection of linkage disequilibrium in *Trypanosoma brucei* isolated from tsetse flies and characterized by RAPD analysis and isoenzymes. *Parasitology* **110**: 181–186.

Waples RS, Do C (2008). LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour* **8**: 753–756.

Watling L, Auster PJ (2005). Distribution of deep-water Alcyonacea off the northeast coast of the United States. In: Freiwald A, Roberts JM (eds). *Cold-Water Corals and Ecosystems*. Springer-Verlag: Berlin, Heidelberg, Germany, pp 279–296.

Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

Wielgoss S, Taraschewski H, Meyer A, Wirth T (2008). Population structure of the parasitic nematode *Anguillicola crassus*, an invader of declining North Atlantic eel stocks. *Mol Ecol* **17**: 3478–3495.

Wilson GA, Rannala B (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**: 1177–1191.

Wolfram K, Mark FC, John U, Lucassen M, Pörtner HO (2006). Microsatellite DNA variation indicates low levels of genetic differentiation among cuttlefish (*Sepia officinalis*) populations in the English Channel and the Bay of Biscay. *Comp Biochem Physiol D* **1**: 375–383.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)

Conservation Genetics Resources
https://doi.org/10.1007/s12686-018-1001-8

TECHNICAL NOTE

SNP discovery in European lobster (*Homarus gammarus*) using RAD sequencing

Tom L. Jenkins[1] · Charlie D. Ellis[2] · Jamie R. Stevens[1]

Received: 17 January 2018 / Accepted: 20 January 2018

Abstract
The European lobster (*Homarus gammarus*) is a decapod crustacean with a high market value and therefore their fisheries are of major importance to the economies they support. However, over-exploitation has led to profound stock declines in some regions such as Scandinavia and the Mediterranean. To manage this resource sustainably, knowledge of population structure and connectivity is crucial to inform management about dispersal, recruitment, stock identification and food traceability. We used restriction-site associated DNA sequencing to develop novel SNP markers from 55 individuals encompassing much of the species range; SNPs were quality filtered, ranked using *F*-statistics and the top 96 SNPs adequate for primer design were retained. SNP markers were developed with the aim of maximising the power to detect genetic differentiation between: (i) Atlantic and Mediterranean lobsters and (ii) Atlantic lobsters. This panel of SNPs provides a useful resource for future studies of population genetic structure and assignment in *H. gammarus*.

Keywords Conservation genetics · Fisheries management · *Homarus gammarus* · Population assignment · RAD-seq · Single nucleotide polymorphism

The European lobster (*Homarus gammarus*) is a decapod crustacean belonging to the family Nephropidae. They are found on hard substrates hiding in crevices or on compressed muds, typically at depths from the low tide mark to 50 m, but they can occur at depths up to 150 m. *Homarus gammarus* is widely distributed, ranging from Morocco to Arctic Norway, including Skagerrak, and also in the Mediterranean where they are generally found more sparsely. The species' high market value makes it a highly-prized seafood product, so its fisheries are of great importance to the local and regional economies they support. However, current and historical over-exploitation has led to stock declines, some of which have been quite profound in several regions (e.g. Scandinavia, Mediterranean) and from which recovery has been slow or stagnant (Kleiven et al. 2012). This has led to the rearing of *H. gammarus* larvae in lobster hatcheries to produce juveniles which are released into the wild to supplement productive stocks where the risk of over-exploitation is high (Ellis et al. 2015).

Over the last decade, genetic diversity and population structure has been investigated in *H. gammarus* using traditional molecular markers including random amplification of polymorphic DNA (RAPDs) (Ulrich et al. 2001), allozymes (Jorstad et al. 2005), mtDNA restriction fragment length polymorphisms (RFLPs) (Triantafyllidis et al. 2005) and microsatellites (Huserbraten et al. 2013; Watson et al. 2016; Ellis et al. 2017). However, single nucleotide polymorphisms (SNPs) are becoming the marker of choice in molecular ecology studies, particularly for non-model organisms without a well-annotated genome, because they are (i) abundant and generally widespread in the genome, (ii) eligible for high-throughput screening and automation, and (iii) reproducible across labs (Seeb et al. 2011). Moreover, genomics now enables thousands to tens of thousands of SNPs to be discovered in non-model marine organisms, meaning we have greater power over previous genetic markers to resolve spatial patterns of genetic differentiation, which is thought

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12686-018-1001-8) contains supplementary material, which is available to authorized users.

✉ Tom L. Jenkins
t.l.jenkins@exeter.ac.uk

[1] Department of Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK

[2] National Lobster Hatchery, South Quay, Padstow, Cornwall PL28 8BL, UK

Published online: 31 January 2018

Table 1 Summary information for the 96 SNP markers developed for the European lobster (*Homarus gammarus*)

Locus ID	Sequence length (bp)	SNP	H_o	H_e	MAF	F_{IS}	P_{HWE}
H_gam_03441	442	G/A	0.407	0.460	0.355	0.045	0.640
H_gam_04173	496	C/T	0.599	0.492	0.409	−0.457	0.273
H_gam_06157	264	G/C	0.383	0.425	0.282	−0.094	0.701
H_gam_07502	97	C/T	0.568	0.499	0.445	−0.184	0.574
H_gam_07892	97	A/T	0.204	0.268	0.155	0.089	0.273
H_gam_08953	496	G/T	0.222	0.423	0.308	0.470	0.018
H_gam_09441	496	A/G	0.414	0.378	0.264	−0.080	0.691
H_gam_11071	400	G/A	0.179	0.251	0.145	0.304	0.239
H_gam_11183	130	A/G	0.537	0.496	0.445	−0.072	0.716
H_gam_11291	270	T/G	0.167	0.213	0.120	−0.274	0.306
H_gam_12971	496	A/G	0.395	0.426	0.309	0.056	0.702
H_gam_14047	496	C/T	0.401	0.417	0.300	−0.252	0.759
H_gam_14742	496	G/A	0.216	0.331	0.222	0.217	0.097
H_gam_15109	496	T/A	0.265	0.423	0.300	0.215	0.087
H_gam_15128	142	C/T	0.383	0.425	0.291	−0.094	0.532
H_gam_15435	496	C/T	0.173	0.190	0.109	0.029	0.611
H_gam_15531	122	G/A	0.284	0.476	0.391	0.074	0.029
H_gam_15581	107	A/G	0.290	0.365	0.236	0.013	0.298
H_gam_18512	496	G/T	0.290	0.337	0.218	−0.179	0.426
H_gam_18652	201	A/G	0.451	0.473	0.364	0.037	0.712
H_gam_19266	175	C/T	0.284	0.296	0.182	−0.125	0.759
H_gam_19460	247	C/T	0.432	0.477	0.382	0.016	0.646
H_gam_20354	142	C/T	0.469	0.430	0.309	−0.194	0.626
H_gam_21197	163	C/T	0.525	0.498	0.436	−0.197	0.759
H_gam_21880	496	A/C	0.481	0.503	0.463	0.054	0.706
H_gam_22323	439	G/A	0.586	0.491	0.418	−0.358	0.465
H_gam_22365	176	A/T	0.340	0.449	0.318	−0.038	0.291
H_gam_22740	138	T/C	0.370	0.386	0.245	−0.328	0.689
H_gam_23146	174	T/C	0.315	0.442	0.300	0.156	0.206
H_gam_23447	114	T/C	0.358	0.472	0.382	0.229	0.275
H_gam_23481	137	T/A	0.296	0.369	0.236	0.159	0.307
H_gam_23677	228	A/G	0.370	0.488	0.418	0.202	0.229
H_gam_23787	496	T/G	0.185	0.246	0.155	0.078	0.267
H_gam_24020	496	C/G	0.216	0.223	0.127	−0.076	0.759
H_gam_25229	230	C/G	0.259	0.227	0.118	−0.395	0.759
H_gam_25580	101	C/T	0.630	0.497	0.436	−0.302	0.276
H_gam_25608	97	C/T	0.185	0.264	0.164	0.268	0.161
H_gam_27329	97	T/C	0.407	0.504	0.464	0.122	0.462
H_gam_28357	496	G/A	0.444	0.497	0.436	0.175	0.587
H_gam_29410	97	T/C	0.420	0.476	0.391	0.001	0.553
H_gam_29801	496	A/G	0.179	0.267	0.164	0.100	0.113
H_gam_29889	496	A/G	0.383	0.483	0.400	0.172	0.451
H_gam_30339	496	G/A	0.228	0.231	0.136	−0.285	0.759
H_gam_31462	140	C/A	0.327	0.455	0.345	0.211	0.198
H_gam_31618	496	A/G	0.333	0.369	0.236	0.059	0.595
H_gam_31967	195	A/C	0.302	0.429	0.318	0.203	0.180
H_gam_31979	182	G/T	0.259	0.380	0.245	−0.080	0.223
H_gam_32358	496	G/A	0.074	0.198	0.109	−0.169	0.036
H_gam_32362	213	C/T	0.630	0.497	0.436	−0.302	0.276
H_gam_32435	496	T/C	0.210	0.246	0.145	0.070	0.489

Table 1 (continued)

Locus ID	Sequence length (bp)	SNP	H_o	H_e	MAF	F_{IS}	P_{HWE}
H_gam_33066	218	C/A	0.370	0.386	0.245	−0.328	0.685
H_gam_33784	136	A/G	0.463	0.504	0.491	0.002	0.715
H_gam_34443	302	G/A	0.346	0.453	0.327	0.186	0.215
H_gam_34818	192	A/C	0.259	0.281	0.173	0.066	0.671
H_gam_35584	97	A/T	0.346	0.445	0.336	0.149	0.306
H_gam_36910	97	A/G	0.395	0.482	0.400	0.096	0.458
H_gam_39107	127	C/T	0.216	0.223	0.127	0.016	0.759
H_gam_39876	134	C/T	0.296	0.312	0.200	−0.155	0.574
H_gam_41521	97	A/T	0.438	0.451	0.355	0.051	0.759
H_gam_42395	496	T/C	0.314	0.472	0.380	0.107	0.119
H_gam_42529	496	A/C	0.364	0.365	0.227	−0.166	0.759
H_gam_42821	190	G/A	0.167	0.185	0.100	0.006	0.581
H_gam_44670	251	T/C	0.204	0.398	0.255	0.402	0.000
H_gam_45154	496	G/A	0.377	0.470	0.373	0.207	0.472
H_gam_45217	496	G/A	0.265	0.259	0.145	−0.136	0.759
H_gam_51159	97	T/G	0.432	0.398	0.273	−0.283	0.692
H_gam_51507	97	G/A	0.308	0.357	0.224	−0.250	0.443
H_gam_53052	496	A/T	0.407	0.368	0.227	−0.288	0.684
H_gam_53263	496	T/A	0.383	0.392	0.255	−0.304	0.691
H_gam_53314	496	T/C	0.327	0.345	0.218	−0.114	0.691
H_gam_53720	96	C/T	0.568	0.495	0.435	−0.335	0.483
H_gam_53889	496	G/C	0.191	0.194	0.118	−0.016	0.631
H_gam_53935	468	C/T	0.284	0.476	0.391	0.074	0.018
H_gam_54240	97	A/C	0.444	0.420	0.287	−0.182	0.759
H_gam_54762	496	C/T	0.531	0.491	0.436	−0.345	0.651
H_gam_55111	146	C/T	0.488	0.503	0.500	−0.222	0.759
H_gam_55142	178	T/G	0.327	0.490	0.426	0.270	0.164
H_gam_55564	496	G/A	0.370	0.503	0.482	0.128	0.264
H_gam_56423	182	C/T	0.420	0.427	0.291	0.127	0.705
H_gam_56785	99	T/C	0.444	0.497	0.436	0.175	0.575
H_gam_57131	97	T/G	0.377	0.407	0.282	−0.090	0.698
H_gam_57989	408	A/T	0.451	0.450	0.336	−0.027	0.759
H_gam_58053	97	A/G	0.049	0.179	0.100	0.046	0.000
H_gam_59503	97	T/A	0.593	0.492	0.427	−0.274	0.335
H_gam_59586	201	G/T	0.296	0.487	0.394	0.261	0.062
H_gam_59967	178	C/T	0.358	0.382	0.255	0.090	0.693
H_gam_60546	167	C/A	0.333	0.494	0.427	0.252	0.166
H_gam_63140	496	C/T	0.321	0.341	0.209	−0.070	0.683
H_gam_63267	97	G/C	0.395	0.381	0.255	−0.085	0.759
H_gam_63581	139	T/C	0.426	0.437	0.318	−0.101	0.705
H_gam_63605	132	T/C	0.451	0.487	0.409	−0.147	0.716
H_gam_63771	97	A/G	0.346	0.454	0.343	0.227	0.287
H_gam_63798	188	G/A	0.568	0.486	0.418	−0.237	0.443
H_gam_65064	496	C/A	0.370	0.386	0.245	−0.328	0.685
H_gam_65376	496	C/A	0.364	0.429	0.309	0.134	0.511
H_gam_65576	173	A/C	0.352	0.376	0.236	−0.457	0.592

Sequences and additional SNP information can be found in S4 Supplementary Material

SNP single nucleotide polymorphism, H_o observed heterozygosity, H_e expected heterozygosity, *MAF* minor allele frequency, F_{IS} inbreeding coefficient, P_{HWE} *P*-values for Hardy–Weinberg equilibrium corrected for multiple comparisons using the false discovery rate

to be particularly beneficial when studying highly dispersive marine species that exhibit typically weak genetic differentiation (e.g. American lobster, Benestan et al. 2015). These advances have also led to the development of small panels of informative SNPs (e.g. Nielsen et al. 2012; Villacorta-Rath et al. 2016) that are likely to be useful for assessments of genetic structure, population assignment and connectivity.

In this study, we used restriction-site associated DNA (RAD) sequencing to isolate and characterise 96 novel SNP markers in *H. gammarus*. Genomic DNA was extracted from v-notch or pleopod tissue using a modified salting-out protocol (Li et al. 2011) (S1 Supplementary Material). The RAD library was prepared in-house using Illumina Nextera XT barcodes and comprised 55 individuals from 27 geographically separate sampling locations, ranging from the Mediterranean to the British Isles and Skagerrak (S2 Supplementary Material). The library was sequenced on an Illumina HiSeq 100 bp paired-end rapid run platform. Raw reads (available from Dryad, https://doi.org/10.5061/dryad.2pc6v) were cleaned and truncated to 97 bp using the process_radtags program in Stacks v1.45 (Catchen et al. 2013) and RAD loci were built using the denovo_map.pl wrapper script in Stacks using optimised parameters of $m=3$, $M=3$ and $n=3$ following the methods of Paris et al. (2017). The populations program was run using all 55 individuals and initial results indicated genetic differentiation between Mediterranean, Skagerrak and the remaining Atlantic samples (S3 Supplementary Material). Therefore, the program was also re-run using only samples from the Atlantic (excluding Mediterranean and Skagerrak samples). This approach maximised the potential to find SNPs that are most informative for detecting hierarchical genetic differentiation between Atlantic lobsters. Full details of the bioinformatics and parameters used are available in S3 Supplementary Information.

In total, 276 million reads were generated and a mean average of 97.9% across all samples were retained after quality control. After initial filtering in Stacks, 7022 biallelic SNPs were identified using all samples and 4377 biallelic SNPs were identified using only Atlantic samples. These SNPs were then ranked by highest G''_{ST} (Meirmans and Hedrick 2011), sorted by the number of SNPs per RAD locus, and filtered for primer design adequacy and suitability for high-throughput genotyping on a Fluidigm EP1 system. The SNP panel was composed of the highest-ranked remaining SNPs; 21 SNPs were chosen from the dataset composed of all samples (aiming to capture differentiation between Atlantic and Mediterranean lobsters) and 78 SNPs were chosen from the dataset composed of only Atlantic samples (aiming to capture any potential hierarchical differentiation in the Atlantic).

Using these 96 SNP markers and all of our samples, we calculated several population genetic statistics for each locus (Table 1). The observed and expected heterozygosity ranged from 0.049 to 0.630 and 0.179 to 0.504, respectively. The minor allele frequency and the inbreeding coefficient ranged from 0.100 to 0.504 and −0.457 to 0.470, respectively. After false discovery rate correction, six SNPs deviated significantly from Hardy–Weinberg equilibrium ($P<0.05$). To our knowledge, this is the first development of SNP markers in *H. gammarus*, and therefore these novel markers offer a valuable tool for future studies of spatial genetic structure and population assignment in this species.

Acknowledgements We thank the many people who provided lobster tissue samples for this research. We also thank Karen Moore and staff at the Exeter Sequencing Service (Exeter, UK) for constructing the RAD libraries. This research was funded by a Natural Environment Research Council UK GW4 + DTP studentship (Grant No. NE/L002434/1), Natural England (Ref. PO 904130) and the University of Exeter, and forms part of the PhD of TLJ.

References

Benestan L, Gosselin T, Perrier C et al (2015) RAD-genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species; the American lobster (*Homarus americanus*). Mol Ecol 24:3299–3315. https://doi.org/10.1111/mec.13245

Catchen J, Hohenlohe PA, Bassham S et al (2013) Stacks: an analysis tool set for population genomics. Mol Ecol 22:3124–3140. https://doi.org/10.1111/mec.12354

Ellis CD, Hodgson DJ, Daniels CL et al (2015) European lobster stocking requires comprehensive impact assessment to determine fishery benefits. ICES J Mar Sci 72:i35–i48. https://doi.org/10.1093/icesjms/fsu196

Ellis CD, Hodgson DJ, Daniels CL et al (2017) Population genetic structure in European lobsters: implications for connectivity, diversity and hatchery stocking. Mar Ecol Prog Ser 563:123–137. https://doi.org/10.3354/meps11957

Huserbraten MBO, Moland E, Knutsen H et al (2013) Conservation, spillover and gene flow within a network of northern European Marine Protected Areas. PLoS ONE 8:e73388. https://doi.org/10.1371/journal.pone.0073388

Jorstad KE, Faresteit E, Kelly E, Triantaphyllids C (2005) Allozyme variation in European lobster (*Homarus gammarus*) throughout its distribution range. N Z J Mar Freshw Res 39:515–526. https://doi.org/10.1080/00288330.2005.9517330

Kleiven AR, Olsen EM, Vølstad JH (2012) Total catch of a red-listed marine species is an order of magnitude higher than official data. PLoS ONE 7:1–7. https://doi.org/10.1371/journal.pone.0031216

Li Y, Wang W, Liu X et al (2011) DNA extraction from crayfish exoskeleton. Indian J Exp Biol 49:953–957

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. Mol Ecol Resour 11:5–18. https://doi.org/10.1111/j.1755-0998.2010.02927.x

Nielsen EE, Cariani A, Aoidh EM et al (2012) Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. Nat Commun 3:851. https://doi.org/10.1038/ncomms1845

Paris JR, Stevens JR, Catchen JM (2017) Lost in parameter space: a road map for stacks. Methods Ecol Evol 8:1360–1373. https://doi.org/10.1111/2041-210X.12775

Seeb JE, Carvalho G, Hauser L et al (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Mol Ecol Resour 11:1–8. https://doi.org/10.1111/j.1755-0998.2010.02979.x

Triantafyllidis A, Apostolidis AP, Katsares V et al (2005) Mitochondrial DNA variation in the European lobster (*Homarus gammarus*) throughout the range. Mar Biol 146:223–235. https://doi.org/10.1007/s00227-004-1435-2

Ulrich I, Muller J, Schutt C, Buchholz F (2001) A study of population genetics in the European lobster, *Homarus gammarus* (Decapoda, Nephropidae). Crustaceana 74:825–837. https://doi.org/10.1163/15685400152682593

Villacorta-Rath C, Ilyushkina I, Strugnell JM et al (2016) Outlier SNPs enable food traceability of the southern rock lobster, *Jasus edwardsii*. Mar Biol 163:163:223. https://doi.org/10.1007/s00227-016-3000-1

Watson HV, McKeown NJ, Coscia I et al (2016) Population genetic structure of the European lobster (*Homarus gammarus*) in the Irish Sea and implications for the effectiveness of the first British marine protected area. Fish Res 183:287–293. https://doi.org/10.1016/j.fishres.2016.06.015

References

Adams, T.P., Miller, R.G., Aleynik, D. & Burrows, M.T. (2014) Offshore marine renewable energy devices as stepping stones across biogeographical boundaries. *Journal of Applied Ecology* **51**, 330–338.

Agnalt, A.L. (2008) Fecundity of the European lobster (*Homarus gammarus*) off southwestern Norway after stock enhancement: Do cultured females produce as many eggs as wild females? *ICES Journal of Marine Science* **65**, 164–170.

Agnalt, A.L., Jorstad, K.E., Kristiansen, T., Nostvold, E., Farestveit, E., Naess, H., Paulsen, O.I. & Svasand, T. (2004) Enhancing the European lobster (*Homarus gammarus*) stock at Kvitsoy Islands: perspectives on rebuilding Norwegian stocks. *Stock Enhancement and Sea Ranching Developments, Pitfalls and Opportunities* pp. 415–426.

Alberto, F., Massa, S., Manent, P., Diaz-Almela, E., Arnaud-Haond, S., Duarte, C.M. & Serrão, E.A. (2008) Genetic differentiation and secondary contact zone in the seagrass *Cymodocea nodosa* across the Mediterranean-Atlantic transition region. *Journal of Biogeography* **35**, 1279–1294.

Allendorf, F.W. (2017) Genetics and the conservation of natural populations: allozymes to genomes. *Molecular Ecology* **26**, 420–430.

Allendorf, F.W., Hohenlohe, P.A. & Luikart, G. (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics* **11**, 697–709.

Almada, F., Francisco, S.M., Lima, C.S., FitzGerald, R., Mirimin, L., Villegas-Ríos, D., Saborido-Rey, F., Afonso, P., Morato, T., Bexiga, S. & Robalo, J.I. (2017) Historical gene flow constraints in a northeastern Atlantic fish: phylogeography of the ballan wrasse *Labrus bergylta* across its distribution range. *Royal Society Open Science* **4**, 160773.

Almany, G.R., Connolly, S.R., Heath, D.D., Hogan, J.D., Jones, G.P., McCook, L.J., Mills, M., Pressey, R.L. & Williamson, D.H. (2009) Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs. *Coral Reefs* **28**, 339–351.

Alvarado-Bremer, J.R., Vinas, J., Mejuto, J., Ely, B. & Pla, C. (2005) Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. *Molecular Phylogenetics and Evolution* **36**, 169–187.

Álvarez-Fernández, I., Fernández, N., Sánchez-Carnero, N. & Freire, J. (2016) The management performance of marine protected areas in the northeast Atlantic Ocean. *Marine Policy* **76**, 159–168.

Anderson, E.C. (2010) Assessing the power of informative subsets of loci for population assignment: Standard methods are upwardly biased. *Molecular Ecology Resources* **10**, 701–710.

Andrews, K. & Luikart, G. (2014) Recent novel approaches for population genomics data analysis. *Molecular Ecology* **23**, 1661–1667.

Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G. & Hohenlohe, P.A. (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* **17**, 81–92.

Araki, H., Cooper, B. & Blouin, M.S. (2007) Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science* **318**, 100–103.

Ardron, J.A. (2008a) The challenge of assessing whether the OSPAR network of marine protected areas is ecologically coherent. *Hydrobiologia* **606**, 45–53.

Ardron, J.A. (2008b) Three initial OSPAR tests of ecological coherence: heuristics in a data-limited situation. *ICES Journal of Marine Science* **65**, 1527–1533.

Aurelle, D., Pivotto, I.D., Malfant, M., Topçu, N.E., Masmoudi, M.B., Chaoui, L., Kara, H.M., Coelho, M.A., Castilho, R. & Haguenauer, A. (2017) Fuzzy species limits in Mediterranean gorgonians (Cnidaria, Octocorallia): inferences on speciation processes. *Zoologica Scripta* **46**, 767–778.

Avise, J.C. (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography* **36**, 3–15.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. & Saunders, N.C. (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**, 489–522.

Baerwald, M., Stephens, M., Bork, K., Meek, M., Tomalty, K. & May, B. (2011) Spring-run chinook salmon genetic management plan - San Joaquin River restoration program. Tech. rep., University of California.

Baetscher, D.S., Hasselman, D.J., Reid, K., Palkovacs, E.P. & Garza, J.C. (2017) Discovery and characterization of single nucleotide polymorphisms in two anadromous alosine fishes of conservation concern. *Ecology and Evolution* **7**, 6638–6648.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. & Johnson, E.A. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* **3**, e3376.

Bannister, R.C. & Addison, J.T. (1998) Enhancing lobster stocks: a review of recent European methods, results, and future prospects. *Bulletin of Marine Science* **62**, 369–387.

Barbazuk, W.B. & Schnable, P.S. (2011) SNP discovery by transcriptome pyrosequencing. *Methods in Molecular Biology* **729**, 225–246.

Barbosa, S., Mestre, F., White, T.A., Paupério, J., Alves, P.C. & Searle, J.B. (2018) Integrative approaches to guide conservation decisions: using genomics to define conservation units and functional corridors. *Molecular Ecology* **27**, 3452–3465.

Bargelloni, L., Alarcon, J.A., Alvarez, M.C., Penzo, E., Magoulas, A., Palma, J. & Patarnello, T. (2005) The Atlantic-Mediterranean transition: discordant genetic patterns in two seabream species, *Diplodus puntazzo* and *Diplodus sargus*. *Molecular Phylogenetics and Evolution* **36**, 523–535.

Bargelloni, L., Alarcon, J.A., Alvarez, M.C., Penzo, E., Magoulas, A., Reis, C. & Patarnello, T. (2003) Discord in the family Sparidae (Teleostei): divergent phylogeographical patterns across the Atlantic-Mediterranean divide. *Journal of Evolutionary Biology* **16**, 1149–1158.

Barido-Sottani, J., Bošková, V., Plessis, L.D., Kühnert, D., Magnus, C., Mitov, V., Müller, N.F., Pečerska, J., Rasmussen, D.A., Zhang, C., Drummond, A.J., Heath, T.A., Pybus, O.G., Vaughan, T.G. & Stadler, T. (2018) Taming the BEAST - a community teaching material resource for BEAST 2. *Systematic Biology* **67**, 170–174.

Baus, E., Darrock, D.J. & Bruford, M.W. (2005) Gene-flow patterns in Atlantic and Mediterranean populations of the Lusitanian sea star *Asterina gibbosa*. *Molecular Ecology* **14**, 3373–3382.

Beal, B.F., Mercer, J.P. & O'Conghaile, A. (2002) Survival and growth of hatchery-reared individuals of the European lobster, *Homarus gammarus* (L.), in field-based nursery cages on the Irish west coat. *Aquaculture* **210**, 137–157.

Becquet, V., Simon-Bouhet, B., Pante, E., Hummel, H. & Garcia, P. (2012) Glacial refugium versus range limit: conservation genetics of *Macoma balthica*, a key species in the Bay of Biscay (France). *Journal of Experimental Marine Biology and Ecology* **432-433**, 73–82.

Beger, M., Selkoe, K., Treml, E., Barber, P., Heyden, v.d.S., Crandall, E., Toonen, R. & Riginos, C. (2014) Evolving coral reef conservation with genetic information. *Bulletin of Marine Science* **90**, 159–185.

Bell, J.D., Leber, K.M., Blankenship, H.L., Loneragan, N.R. & Masuda, R. (2008) A new era for restocking, stock enhancement and sea ranching of coastal fisheries resources. *Reviews in Fisheries Science* **16**, 1–9.

Benavente, G.P., Uglem, I., Browne, R. & Balsa, C.M. (2010) Culture of juvenile European lobster (*Homarus gammarus* L.) in submerged cages. *Aquaculture International* **18**, 1177–1189.

Benestan, L., Ferchaud, A.L., Hohenlohe, P., Garner, B.A., Naylor, G.J., Baums, I., Schwartz, M., Kelley, J.L. & Luikart, G. (2016a) Conservation genomics of natural and managed populations: building a conceptual and practical framework. *Molecular Ecology* **25**, 2967–2977.

Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R. & Bernatchez, L. (2015) RAD-genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species; the American lobster (*Homarus americanus*). *Molecular Ecology* **24**, 3299–3315.

Benestan, L., Quinn, B.K., Maaroufi, H., Laporte, M., Rochette, R. & Bernatchez, L. (2016b) Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology* **25**, 5073–5092.

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D.T., Barth, J.M., Beacham, T.D., Maes, G.E., Martinsohn, J.T., Miller, K.M., Naish, K.A., Ovenden, J.R., Primmer, C.R., Young Suk, H., Therkildsen, N.O. & Withler, R.E. (2017) Harnessing the power of genomics to secure the future of seafood. *Trends in Ecology & Evolution* **32**, 665–680.

Bhat, S., Polanowski, A.M., Double, M.C., Jarman, S.N. & Emslie, K.R. (2012) The effect of input DNA copy number on genotype call and characterising SNP markers in the humpback whale genome using a nanofluidic array. *PLoS ONE* **7**.

Bierne, N., Gagnaire, P.A. & David, P. (2013) The geography of introgression in a patchy environment and the thorn in the side of ecological speciation. *Current Zoology* **59**, 72–86.

Bishop, M.J., Mayer-Pinto, M., Airoldi, L., Firth, L.B., Morris, R.L., Loke, L.H., Hawkins, S.J., Naylor, L.A., Coleman, R.A., Chee, S.Y. & Dafforn, K.A. (2017) Effects of ocean sprawl on ecological connectivity: impacts and solutions **492**, 7–30.

Borcard, D. & Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling* **153**, 51–68.

Botsford, L.W., White, J.W., Coffroth, M.A., Paris, C.B., Planes, S., Shearer, T.L., Thorrold, S.R. & Jones, G.P. (2009) Connectivity and resilience of coral reef metapopulations in marine protected areas: matching empirical efforts to predictive needs. *Coral Reefs* **28**, 327–337.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10**, e1003537.

Bradbury, I.R., Laurel, B., Snelgrove, P.V., Bentzen, P. & Campana, S.E. (2008) Global patterns in marine dispersal estimates: the influence of geography, taxonomic category and life history. *Proceedings of the Royal Society B: Biological Sciences* **275**, 1803–1809.

Breusing, C., Biastoch, A., Drews, A., Metaxas, A., Jollivet, D., Vrijenhoek, R.C., Bayer, T., Melzner, F., Sayavedra, L., Petersen, J.M., Dubilier, N., Schilhabel, M.B., Rosenstiel, P. & Reusch, T.B.H. (2016) Biophysical and population genetic models predict the presence of "phantom" stepping stones connecting Mid-Atlantic ridge vent ecosystems. *Current Biology* **26**, 2257–2267.

Britt, M., Haworth, S.E., Johnson, J.B., Martchenko, D. & Shafer, A.B. (2018) The importance of non-academic coauthors in bridging the conservation genetics gap. *Biological Conservation* **218**, 118–123.

Brown, J.H. & Kodric-Brown, A. (1977) Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology* **58**, 445–449.

Calvignac, S., Konecny, L., Malard, F. & Douady, C.J. (2011) Preventing the pollution of mitochondrial datasets with nuclear mitochondrial paralogs (numts). *Mitochondrion* **11**, 246–254.

Campbell, E.O., Brunet, B.M., Dupuis, J.R. & Sperling, F.A. (2018) Would an RRS by any other name sound as RAD? *Methods in Ecology and Evolution* **9**, 1920–1927.

Carr, H., Cornthwaite, A., Wright, H. & Davies, J. (2014) Assessing progress towards an ecologically coherent network of MPAs in Secretary of State Waters in 2014. Tech. rep., JNCC, JNCC Report.

Carr, H., Wright, H., Cornthwaite, A. & Davies, J. (2016) Assessing the contribution of Welsh MPAs towards an ecologically coherent MPA network in 2016. Tech. rep., JNCC, JNCC Report.

Casale, P., Laurent, L., Gerosa, G. & Argano, R. (2002) Molecular evidence of male-biased dispersal in loggerhead turtle juveniles. *Journal of Experimental Marine Biology and Ecology* **267**, 139–145.

Castilho, R., Cunha, R.L., Faria, C., Velasco, E.M. & Robalo, J.I. (2017) Asymmetrical dispersal and putative isolation-by-distance of an intertidal blenniid across the Atlantic-Mediterranean divide. *PeerJ* **5**, e3195.

Casu, M., Sanna, D., Cossu, P., Lai, T., Francalacci, P. & Curini-Galletti, M. (2011) Molecular phylogeography of the microturbellarian *Monocelis lineata* (Platyhelminthes: Proseriata) in the North-East Atlantic. *Biological Journal of the Linnean Society* **103**, 117–135.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A. & Cresko, W.A. (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology* **22**, 3124–3140.

Cayuela, H., Rougemont, Q., Prunier, J.G., Moore, J.S., Clobert, J., Besnard, A. & Bernatchez, L. (2018) Demographic and genetic approaches to study dispersal in wild animal populations: a methodological review. *Molecular Ecology* **27**, 3976–4010.

CEFAS (2017) Lobster (*Homarus gammarus*) CEFAS stock status report 2017. Tech. rep.

Chambers, J.M. (2013) *SoDA: functions and examples for "Software for Data Analysis"*. R package version 1.0-6.

Charlesworth, B. (2009) Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics* **10**, 195–205.

Charrier, G., Chenel, T., Durand, J.D., Girard, M., Quiniou, L. & Laroche, J. (2006) Discrepancies in phylogeographical patterns of two European anglerfishes (*Lophius budegassa* and *Lophius piscatorius*). *Molecular Phylogenetics and Evolution* **38**, 742–754.

Chattopadhyay, B., Garg, K.M. & Ramakrishnan, U. (2014) Effect of diversity and missing data on genetic assignment with RAD-Seq markers. *BMC Research Notes* **7**, 4–6.

Chen, K.Y., Marschall, E.A., Sovic, M.G., Fries, A.C., Gibbs, H.L. & Ludsin, S.A. (2018) assignPOP: An r package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. *Methods in Ecology and Evolution* **9**, 439–446.

Chevolot, M., Hoarau, G., Rijnsdorp, A.D., Stam, W.T. & Olsen, J.L. (2006) Phylogeography and population structure of thornback rays (*Raja clavata* L., Rajidae). *Molecular Ecology* **15**, 3693–3705.

Chhatre, V.E. & Emerson, K.J. (2017) StrAuto: Automation and parallelization of STRUCTURE analysis. *BMC Bioinformatics* **18**, 192.

Chiverrell, R.C. & Thomas, G.S.P. (2010) Extent and timing of the Last Glacial Maximum (LGM) in Britain and Ireland: a review. *Journal of Quaternary Science* **25**, 535–549.

Christie, M.R., Meirmans, P.G., Gaggiotti, O.E., Toonen, R.J. & White, C. (2017) Disentangling the relative merits and disadvantages of parentage analysis and assignment tests for inferring population connectivity. *ICES Journal of Marine Science* **30**, 913–924.

Christie, M.R., Tissot, B.N., Albins, M.A., Beets, J.P., Jia, Y., Ortiz, D.M., Thompson, S.E. & Hixon, M.A. (2010) Larval connectivity in an effective network of marine protected areas. *PLoS ONE* **5**, e15715.

Clark, P., Dyke, A., Shakun, J., Carlson, A., Clark, J., Wohlfarth, B., Mitrovica, J., Hostetler, S. & McCabe, M. (2009) The Last Glacial Maximum. *Science* **325**, 710–714.

Coleman, M.A., Chambers, J., Knott, N.A., Malcolm, H.A., Harasti, D., Jordan, A. & Kelaher, B.P. (2011) Connectivity within and among a network of temperate marine reserves. *PLoS ONE* **6**, e20168.

Cooke, S.J., Midwood, J.D., Thiem, J.D., Klimley, P., Lucas, M.C., Thorstad, E.B., Eiler, J., Holbrook, C. & Ebner, B.C. (2013) Tracking animals in freshwater with electronic tags: past, present and future. *Animal Biotelemetry* **1**, 5.

Correia, A.T., Ramos, A.A., Barros, F., Silva, G., Hamer, P., Morais, P., Cunha, R.L. & Castilho, R. (2012) Population structure and connectivity of the European conger eel (*Conger conger*) across the north-eastern Atlantic and western Mediterranean: integrating molecular and otolith elemental approaches. *Marine Biology* **159**, 1509–1525.

Coscia, I. & Mariani, S. (2011) Phylogeography and population structure of European sea bass in the north-east Atlantic. *Biological Journal of the Linnean Society* **104**, 364–377.

Costa, D.P., Breed, G.A. & Robinson, P.W. (2012) New insights into pelagic migrations: implications for ecology and conservation. *Annual Review of Ecology, Evolution, and Systematics* **43**, 73–96.

Costantini, F., Gori, A., Lopez-González, P., Bramanti, L., Rossi, S., Gili, J.M. & Abbiati, M. (2016) Limited genetic connectivity between gorgonian morphotypes along a depth gradient. *Plos One* **11**, e0160678.

Couceiro, L., Barreiro, R., Ruiz, J.M. & Sotka, E.E. (2007) Genetic isolation by distance among populations of the netted dog whelk *Nassarius reticulatus* (L.) along the European Atlantic coastline. *The Journal of Heredity* **98**, 603–10.

Couceiro, L., López, L., Sotka, E., Ruiz, J. & Barreiro, R. (2012) Molecular data delineate cryptic *Nassarius* species and characterize spatial genetic structure of *N. nitidus*. *Journal of the Marine Biological Association of the United Kingdom* **92**, 1175–1182.

Cowen, R., Gawarkiewicz, G., Pineda, J., Thorrold, S.R. & Werner, F. (2007) Population connectivity in marine systems. *Oceanography* **20**, 14–21.

Cowen, R.K. (2006) Scaling of Connectivity in Marine Populations. *Science* **311**, 522–527.

Cowen, R.K., Lwiza, K.M., Sponaugle, S., Paris, C.B. & Olson, D.B. (2000) Connectivity of marine populations: open or closed? *Science* **287**, 857–859.

Cowen, R.K. & Sponaugle, S. (2009) Larval dispersal and marine population connectivity. *Annual Review of Marine Science* **1**, 443–466.

Coyer, J.A., Peters, A.F., Stam, W.T. & Olsen, J.L. (2003) Post-ice age recolonization and differentiation of *Fucus serratus* L. (Phaeophyceae; Fucaceae) populations in Northern Europe. *Molecular Ecology* **12**, 1817–1829.

Crandall, E.D., Sbrocco, E.J., DeBoer, T.S., Barber, P.H. & Carpenter, K.E. (2012) Expansion dating: calibrating molecular clocks in marine species from expansions onto the Sunda Shelf following the Last Glacial Maximum. *Molecular Biology and Evolution* **29**, 707–719.

Cuéllar-Pinzón, J., Presa, P., Hawkins, S.J. & Pita, A. (2016) Genetic markers in marine fisheries: types, tasks and trends. *Fisheries Research* **173**, 194–205.

Cure, K., Thomas, L., Hobbs, J.P.A., Fairclough, D.V. & Kennington, W.J. (2017) Genomic signatures of local adaptation reveal source-sink dynamics in a high gene flow fish species. *Scientific Reports* **7**, 8618.

Cuveliers, E.L., Larmuseau, M.H.D., Hellemans, B., Verherstraeten, S.L.N.A., Volckaert, F.A.M. & Maes, G.E. (2012) Multi-marker estimate of genetic connectivity of sole (*Solea solea*) in the North-East Atlantic Ocean. *Marine Biology* **159**, 1239–1253.

Da Silva, A.G., Appleyard, S.A. & Upston, J. (2015) Establishing the evolutionary compatibility of potential sources of colonizers for overfished stocks: a population genomics approach. *Molecular Ecology* **24**, 564–579.

Dalongeville, A., Benestan, L., Mouillot, D., Lobreaux, S. & Manel, S. (2018) Combining six genome scan methods to detect candidate genes to salinity in the Mediterranean striped red mullet (*Mullus surmuletus*). *BMC Genomics* **19**, 217.

Daniels, C.L., Wills, B., Ruiz-Perez, M., Miles, E., Wilson, R.W. & Boothroyd, D. (2015) Development of sea based container culture for rearing European lobster (*Homarus gammarus*) around South West England. *Aquaculture* **448**, 186–195.

D'Arcy, J., Mirimin, L. & FitzGerald, R. (2013) Phylogeographic structure of a protogynous hermaphrodite species, the ballan wrasse *Labrus bergylta*, in Ireland, Scotland, and Norway, using mitochondrial DNA. *ICES Journal of Marine Science* **70**, 685–693.

Dauvin, J.C. (2012) Are the eastern and western basins of the English Channel two separate ecosystems? *Marine Pollution Bulletin* **64**, 463–471.

Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M. & Blaxter, M.L. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* **12**, 499–510.

Davies, T.E., Maxwell, S.M., Kaschner, K., Garilao, C. & Ban, N.C. (2017) Large marine protected areas represent biodiversity now and under climate change. *Scientific Reports* **7**, 9569.

DeAngelis, D.L. & Grimm, V. (2014) Individual-based models in ecology after four decades. *F1000Prime Reports* **6**, 39.

Dehens, L.A. & Fanning, L.M. (2018) What counts in making marine protected areas (MPAs) count? The role of legitimacy in MPA success in Canada. *Ecological Indicators* **86**, 45–57.

Devillers, R., Pressey, R.L., Grech, A., Kittinger, J.N., Edgar, G.J., Ward, T. & Watson, R. (2015) Reinventing residual reserves in the sea: are we favouring ease of establishment over need for protection? *Aquatic Conservation: Marine and Freshwater Ecosystems* **25**, 480–504.

Diopere, E., Vandamme, S.G., Hablu, P.I., Cariani, A., Houdt, J.V., Rijnsdorp, A., Tinti, F., Consortium, F., Volckaert, F.A.M. & Maes, G.E. (2017) Seascape genetics of a flatfish reveals local selection under high levels of gene flow. *ICES Journal of Marine Science* **75**, 675–689.

Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., Jombart, T., Larocque, G., Legendre, P., Madi, N. & Wagner, H.H. (2018) *adespatial: Multivariate Multiscale Spatial Analysis*. R package version 0.2-0.

Dray, S. & Dufour, A.B. (2007) *The ade4 Package: implementing the duality diagram for ecologists*. R package version 1.7.11.

Drinan, D.P., Gruenthal, K.M., Canino, M.F., Lowry, D., Fisher, M.C., Hauser, L., Sciences, F. & Fisheries, A. (2018) Population assignment and local adaptation along an isolation-by distance gradient in Pacific cod (*Gadus macrocephalus*). *Evolutionary Applications* **11**, 1448–1464.

Drummond, A.J., Rambaut, A., Shapiro, B. & Pybus, O.G. (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* **22**, 1185–1192.

Drury, C., Schopmeyer, S., Goergen, E., Bartels, E., Nedimyer, K., Johnson, M., Maxwell, K., Galvan, V., Manfrino, C. & Lirman, D. (2017) Genomic patterns in *Acropora cervicornis* show extensive population structure and variable genetic diversity. *Ecology and Evolution* **7**, 6188–6200.

Dufresne, F., Stift, M., Vergilino, R. & Mable, B.K. (2014) Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. *Molecular Ecology* **23**, 40–69.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797.

Ellegren, H. (2008) Sequencing goes 454 and takes large-scale genomics into the wild **17**, 1629–1631.

Ellegren, H. (2014) Genome sequencing and population genomics in non-model organisms. *Trends in Ecology and Evolution* **29**, 51–63.

Ellis, C., Hodgson, D., Daniels, C., Collins, M. & Griffiths, A. (2017) Population genetic structure in European lobsters: implications for connectivity, diversity and hatchery stocking. *Marine Ecology Progress Series* **563**, 123–137.

Ellis, C., Knott, H., Daniels, C., Witt, M. & Hodgson, D. (2015a) Geographic and environmental drivers of fecundity in the European lobster (*Homarus gammarus*). *ICES Journal of Marine Science* **72**, i91–i100.

Ellis, C.D., Hodgson, D.J., André, C., Sørdalen, T.K., Knutsen, H. & Griffiths, A.G.F. (2015b) Genotype reconstruction of paternity in European lobsters (*Homarus gammarus*). *PLoS ONE* **10**, e0139585.

Ellis, C.D., Hodgson, D.J., Daniels, C.L., Boothroyd, D.P., Bannister, R.C.A. & Griffiths, A.G.F. (2015c) European lobster stocking requires comprehensive impact assessment to determine fishery benefits. *ICES Journal of Marine Science* **72**, i35–i48.

Ellis, J.S., Gilbey, J., Armstrong, A., Balstad, T., Cauwelier, E., Cherbonnel, C., Consuegra, S., Coughlan, J., Cross, T.F., Crozier, W., Dillane, E., Ensing, D., de Leániz, C.G., García-Vázquez, E., Griffiths, A.M., Hindar, K., Hjorleifsdottir, S., Knox, D., Machado-Schiaffino, G., McGinnity, P., Meldrup, D., Nielsen, E.E., Olafsson, K., Primmer, C.R., Prodohl, P., Stradmeyer, L., Vähä, J.P., Verspoor, E., Wennevik, V. & Stevens, J.R. (2011) Microsatellite standardization and evaluation of genotyping error in a large multi-partner research programme for conservation of Atlantic salmon (*Salmo salar*). *Genetica* **139**, 353–367.

Emerson, K.J., Merz, C.R., Catchen, J.M., Hohenlohe, P.a., Cresko, W.a., Bradshaw, W.E. & Holzapfel, C.M. (2010) Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences* **107**, 16196–16200.

Etter, P.D., Bassham, S., Hohenlohe, P.A., Johnson, E.A. & Cresko, W.A. (2011) SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods in Molecular Biology* **772**, 157–178.

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* **14**, 2611–2620.

Everett, M.V., Park, L.K., Berntson, E.A., Elz, A.E., Whitmire, C.E., Keller, A.A. & Clarke, M.E. (2016) Large-scale genotyping-by-sequencing indicates high levels of gene flow in the deep-sea octocoral *Swiftia simplex* (Nutting 1909) on the west coast of the United States. *Plos One* **11**, e0165279.

Excoffier, L. & Lischer, H.E.L. (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564–567.

FAO (2018) Food and agriculture organisation of the united nations: *Homarus gammarus*. Tech. rep.

Faust, E., Halvorsen, K.T., Andersen, P., Knutsen, H. & André, C. (2018) Cleaner fish escape salmon farms and hybridize with local wrasse populations. *Royal Society Open Science* **5**, 171752.

Finnegan, A.K., Griffiths, A.M., King, R.A., Machado-Schiaffino, G., Porcher, J.P., Garcia-Vazquez, E., Bright, D. & Stevens, J.R. (2013) Use of multiple markers demonstrates a cryptic western refugium and postglacial colonisation routes of Atlantic salmon (*Salmo salar* L.) in northwest Europe. *Heredity* **111**, 34–43.

Firth, L., Knights, A.M., Bridger, D., Evans, A.J., Mieszkowska, N., Moore, P.J., O'Connor, N.E., Sheehan, E.V., Thompson, R.C. & Hawkins, S.J. (2016) Ocean sprawl: challenges and opportunities for biodiversity management in a changing world. *Oceanography and Marine Biology: An Annual Review* **54**, 189–262.

FIS (2018) Fish information and services: market prices, search by species. Tech. rep.

Flanagan, S.P., Forester, B.R., Latch, E.K., Aitken, S.N. & Hoban, S. (2017) Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. *Evolutionary Applications* **11**, 1035–1052.

Fodrie, F.J., Becker, B.J., Levin, L.A., Gruenthal, K. & McMillan, P.A. (2011) Connectivity clues from short-term variability in settlement and geochemical tags of mytilid mussels. *Journal of Sea Research* **65**, 141–150.

Fodrie, F.J. & Herzka, S.Z. (2013) A comparison of otolith geochemistry and stable isotope markers to track fish movement: describing estuarine ingress by larval and post-larval halibut. *Estuaries and Coasts* **36**, 906–917.

Fogarty, M.J. & Botsford, L.W. (2007) Population connectivity and spatial management of marine fisheries. *Oceanography* **20**, 112–123.

Foll, M. & Gaggiotti, O. (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* **180**, 977–993.

Fontes, J., Caselle, J.E., Sheehy, M.S., Santos, R.S. & Warner, R.R. (2009) Natal signatures of juvenile *Coris julis* in the Azores: investigating connectivity scenarios in an oceanic archipelago. *Marine Ecology Progress Series* **387**, 51–59.

Fossette, S., Katija, K., Goldbogen, J.A., Bograd, S., Patry, W., Howard, M.J., Knowles, T., Haddock, S.H., Bedell, L., Hazen, E.L., Robison, B.H., Mooney, T.A., Shorter, K.A., Bastian, T. & Gleiss, A.C. (2016) How to tag a jellyfish? A methodological review and guidelines to successful jellyfish tagging. *Journal of Plankton Research* **38**, 1347–1363.

Foster, N.L., Rees, S., Langmead, O., Griffiths, C., Oates, J. & Attrill, M.J. (2017) Assessing the ecological coherence of a marine protected area network in the Celtic Seas. *Ecosphere* **8**, e01688.

Francis, R.M. (2017) Pophelper: an r package and web app to analyse and visualize population structure. *Molecular Ecology Resources* **17**, 27–32.

Frankham, R. (2015) Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology* **24**, 2610–2618.

Frankham, R., Ballou, J.D., Eldridge, M.D.B., Lacy, R.C., Ralls, K., Dudash, M.R. & Fenster, C.B. (2011) Predicting the probability of outbreeding depression. *Conservation Biology* **25**, 465–475.

Frantz, A.C., Cellina, S., Krier, A., Schley, L. & Burke, T. (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology* **46**, 493–505.

Fu, Y.X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915–925.

Fu, Z., Epstein, B., Kelley, J.L., Zheng, Q., Bergland, A.O., Castillo Carrillo, C.I., Jensen, A.S., Dahan, J., Karasev, A.V. & Snyder, W.E. (2017) Using NextRAD sequencing to infer movement of herbivores among host plants. *PLoS ONE* **12**, e0177742.

Fuentes-Pardo, A.P. & Ruzzante, D.E. (2017) Whole-genome sequencing approaches for conservation biology: advantages, limitations and practical recommendations. *Molecular Ecology* **26**, 5369–5406.

Funk, W.C., Forester, B.R., Converse, S.J., Darst, C. & Morey, S. (2018) Improving conservation policy with genomics: a guide to integrating adaptive potential into U.S. Endangered Species Act decisions for conservation practitioners and geneticists. *Conservation Genetics* ***in press***, https://doi.org/10.1007/s10592–018–1096–1.

Funk, W.C., McKay, J.K., Hohenlohe, P.A. & Allendorf, F.W. (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution* **27**, 489–496.

Gagnaire, P.A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., Arnaud-Haond, S. & Bierne, N. (2015) Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications* **8**, 769–786.

Galla, S.J., Buckley, T.R., Elshire, R., Hale, M.L., Knapp, M., Mccallum, J., Moraga, R., Santure, A.W., Wilcox, P. & Steeves, T.E. (2016) Building strong relationships between conservation genetics and primary industry leads to mutually beneficial genomic advances. *Molecular Ecology* **25**, 5267–5281.

Galtier, N., Nabholz, B., Glemin, S. & Hurst, G.D.D. (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology* **18**, 4541–4550.

García-Merchán, V.H., Robainas-Barcia, A., Abelló, P., Macpherson, E., Palero, F., García-Rodríguez, M., Gil de Sola, L. & Pascual, M. (2012) Phylogeographic patterns of decapod crustaceans at the Atlantic-Mediterranean transition. *Molecular Phylogenetics and Evolution* **62**, 664–672.

Garner, B.A., Hand, B.K., Amish, S.J., Bernatchez, L., Foster, J.T., Miller, K.M., Morin, P.A., Narum, S.R., O'Brien, S.J., Roffler, G., Templin, W.D., Sunnucks, P., Strait, J., Warheit, K.I., Seamons, T.R., Wenburg, J., Olsen, J. & Luikart, G. (2016) Genomics in conservation: case studies and bridging the gap between data and application. *Trends in Ecology and Evolution* **31**, 81–82.

Gautier, M., Gharbi, K., Cezard, T., Foucaud, J., Kerdelhué, C., Pudlo, P., Cornuet, J.M. & Estoup, A. (2013) The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Molecular Ecology* **22**, 3165–3178.

Gerber, L.R., Del Mark Mancha-Cisneros, M., O'Connor, M. & Selig, E.R. (2014) Climate change impacts on connectivity in the ocean: implications for conservation. *Ecosphere* **5**, 1–18.

Gilbert, K.J. (2016) Identifying the number of population clusters with STRUCTURE: problems and solutions. *Molecular Ecology Resources* **16**, 601–603.

Gill, D.A., Mascia, M.B., Ahmadia, G.N., Glew, L., Lester, S.E., Barnes, M., Craigie, I., Darling, E.S., Free, C.M., Geldmann, J., Holst, S., Jensen, O.P., White, A.T., Basurto, X., Coad, L., Gates, R.D., Guannel, G., Mumby, P.J., Thomas, H., Whitmee, S., Woodley, S. & Fox, H.E. (2017) Capacity shortfalls hinder the performance of marine protected areas globally. *Nature* **543**, 665–669.

Gomes, I., Peteiro, L.G., Albuquerque, R., Nolasco, R., Dubert, J., Swearer, S.E. & Queiroga, H. (2016) Wandering mussels: using natural tags to identify connectivity patterns among Marine Protected Areas. *Marine Ecology Progress Series* **552**, 159–176.

Gómez, A., Hughes, R.N., Wright, P.J., Carvalho, G.R. & Lunt, D.H. (2007) Mitochondrial DNA phylogeography and mating compatibility reveal marked genetic structuring and speciation in the NE Atlantic bryozoan *Celleporella hyalina*. *Molecular Ecology* **16**, 2173–2188.

Gordon, I.J., Evans, D.M., Garner, T.W., Katzner, T., Gompper, M.E., Altwegg, R., Branch, T.A., Johnson, J.A. & Pettorelli, N. (2014) Enhancing communication between conservation biologists and conservation practitioners: letter from the conservation front line. *Animal Conservation* **17**, 1–2.

Gosselin, T. (2017) *radiator: RADseq Data Exploration, Manipulation and Visualization using R*. R package version 0.0.5.

Gosselin, T., Sainte-Marie, B. & Bernatchez, L. (2005) Geographic variation of multiple paternity in the American lobster, *Homarus americanus*. *Molecular Ecology* **14**, 1517–1525.

Graham, C.F., Glenn, T.C., McArthur, A.G., Boreham, D.R., Kieran, T., Lance, S., Manzon, R.G., Martino, J.a., Pierson, T., Rogers, S.M., Wilson, J.Y. & Somers, C.M. (2015) Impacts of degraded DNA on restriction enzyme associated DNA Sequencing (RADSeq). *Molecular Ecology Resources* **15**, 1304–1315.

Green, S.J., Akins, J.L., Maljković, A. & Côté, I.M. (2012) Invasive lionfish drive Atlantic coral reef fish declines. *PLoS ONE* **7**, e32596.

Guillot, G., Jónsson, H., Hinge, A., Manchih, N. & Orlando, L. (2016) Accurate continuous geographic assignment from low- to high-density SNP data. *Bioinformatics* **32**, 1106–1108.

Gysels, E.S., Hellemans, B., Pampoulie, C. & Volckaert, F.A.M. (2004) Phylogeography of the common goby, *Pomatoschistus microps*, with particular emphasis on the colonization of the Mediterranean and the North Sea. *Molecular Ecology* **13**, 403–417.

Habicht, C., Munro, A.R., Dann, T.H., Eggers, D.M., Templin, W.D., Witteveen, M.J., Baker, T.T., Howard, K.G., Jasper, J.R., Rogers Olive, S.D., Liller, H.L., Chenoweth, E.L. & Volk, E.C. (2012) Harvest and harvest rates of sockeye salmon stocks in fisheries of the western Alaska salmon stock identification program (WASSIP). Tech. rep., Alaska Department of Fish and Game, Special Publication No. 12-24, Anchorage, Alaska Department of Fish and Game, Special Publication No. 12-24, Anchorage.

Halanych, K.M., Vodoti, E.T., Sundberg, P. & Dahlgren, T.G. (2013) Phylogeography of the horse mussel *Modiolus modiolus*. *Journal of the Marine Biological Association of the United Kingdom* **93**, 1857–1869.

Hall-Spencer, J.M., Pike, J. & Munn, C.B. (2007) Diseases affect cold-water corals too: *Eunicella verrucosa* (Cnidaria: Gorgonacea) necrosis in SW England. *Diseases of Aquatic Organisms* **76**, 87–97.

Hanski, I. (1998) Metapopulation dynamics. *Nature* **396**, 41–49.

Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R.S., Ruegg, K. & Palstra, F. (2011) Understanding and estimating effective population size for practical application in marine species management. *Conservation Biology* **25**, 438–449.

Harpending, H.C. (1994) Signature of ancient population growth in a low resolution mitochondrial DNA mismatch distribution. *Human Biology* **66**, 591–600.

Harrison, F. (2011) Getting started with meta-analysis. *Methods in Ecology and Evolution* **2**, 1–10.

Harrison, H.B., Williamson, D.H., Evans, R.D., Almany, G.R., Thorrold, S.R., Russ, G.R., Feldheim, K.A., Van Herwerden, L., Planes, S., Srinivasan, M., Berumen, M.L. & Jones, G.P. (2012) Larval export from marine reserves and the recruitment benefit for fish and fisheries. *Current Biology* **22**, 1023–1028.

Harrisson, K.A., Amish, S.J., Pavlova, A., Narum, S.R., Telonis-Scott, M., Rourke, M.L., Lyon, J., Tonkin, Z., Gilligan, D.M., Ingram, B.A., Lintermans, M., Gan, H.M., Austin, C.M., Luikart, G. & Sunnucks, P. (2017) Signatures of polygenic adaptation associated with climate across the range of a threatened fish species with high genetic connectivity. *Molecular Ecology* **26**, 6253–6269.

Hauser, L. & Carvalho, G.R. (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts **9**, 333–362.

Hays, G.C., Houghton, J.D.R. & Myers, A.E. (2004) Endangered species: Pan-Atlantic leatherback turtle movements. *Nature* **429**, 522.

Hayward, P. & Ryland, J. (1995) *Handbook of the marine fauna of north-west Europe*. Oxford University Press, Oxford, United Kingdom.

Hazen, E.L., Maxwell, S.M., Bailey, H., Bograd, S.J., Hamann, M., Gaspar, P., Godley, B.J. & Shillinger, G.L. (2012) Ontogeny in marine tagging and tracking science: technologies and data gaps. *Marine Ecology Progress Series* **457**, 221–240.

Hazkani-Covo, E., Zeller, R.M. & Martin, W. (2010) Molecular poltergeists: Mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genetics* **6**, e1000834.

Hedgecock, D., Barber, P.H. & Edmands, S. (2007) Genetic approaches to measuring connectivity. *Oceanography* **20**, 70–79.

Hellberg, M. & Burton, R. (2002) Genetic assessment of connectivity among marine populations. *Bulletin of Marine Science* **70**, 273–290.

Hellberg, M.E. (2009) Gene flow and isolation among populations of marine animals. *Annual Review of Ecology, Evolution, and Systematics* **40**, 291–310.

Helyar, S.J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M.I., Ogden, R., Limborg, M.T., Cariani, A., Maes, G.E., Diopere, E., Carvalho, G.R. & Nielsen, E.E. (2011) Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources* **11**, 123–136.

Herrera, S. & Shank, T.M. (2016) RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics and Evolution* **100**, 70–79.

Hewitt, G.M. (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68**, 87–112.

Hewitt, G.M. (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907–913.

Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B* **359**, 183–195.

Heyden, S.V.D., Beger, M., Toonen, R.J., Juinio-meñez, M.A., Ravago-gotanco, R., Fauvelot, C. & Bernardi, G. (2014) The application of genetics to marine management. *Bulletin of Marine Sciences* **90**, 123–158.

Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H., Johnson, J.B., Rissler, L., Victoriano, P.F. & Yoder, A.D. (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* **54**, 291–301.

Hidalgo, M., Kaplan, D.M., Kerr, L.A., Watson, J.R., Paris, C.B. & Browman, H.I. (2017) Advancing the link between ocean connectivity, ecological function and management challenges. *ICES Journal of Marine Science* **74**, 1702–1707.

Ho, S.Y.W., Lanfear, R., Bromham, L., Phillips, M.J., Soubrier, J., Rodrigo, A.G. & Cooper, A. (2011) Time-dependent rates of molecular evolution. *Molecular Ecology* **20**, 3087–3101.

Ho, S.Y.W., Saarma, U., Barnett, R., Haile, J. & Shapiro, B. (2008) The effect of inappropriate calibration: three case studies in molecular ecology. *PLoS ONE* **3**, e1615.

Hoarau, G., Coyer, J.A., Veldsink, J.H., Stam, W.T. & Olsen, J.L. (2007) Glacial refugia and recolonization pathways in the brown seaweed *Fucus serratus*. *Molecular Ecology* **16**, 3606–3616.

Hoban, S., Gaggiotti, O. & Bertorelle, G. (2013) Sample planning optimization tool for conservation and population genetics (SPOTG): a software for choosing the appropriate number of markers and samples. *Methods in Ecology and Evolution* **4**, 299–303.

Hodel, R.G., Chen, S., Payton, A.C., McDaniel, S.F., Soltis, P. & Soltis, D.E. (2017) Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: comparing microsatellites and RAD-Seq and investigating loci filtering. *Scientific Reports* **7**, 17598.

Hodges, E., Xuan, Z., Balija, V., Kramer, M., Molla, M.N., Smith, S.W., Middle, C.M., Rodesch, M.J., Albert, T.J., Hannon, G.J. & McCombie, W.R. (2007) Genome-wide in situ exon capture for selective resequencing. *Nature Genetics* **39**, 1522–1527.

Hofreiter, M. & Stewart, J. (2009) Ecological change, range fluctuations and population dynamics during the Pleistocene. *Current Biology* **19**, R584–R594.

Hogg, C.J., Grueber, C.E., Pemberton, D., Fox, S., Lee, A.V., Ivy, J.A. & Belov, K. (2017) "Devil Tools & Tech": a synergy of conservation research and management practice. *Conservation Letters* **10**, 133–138.

Holland, L.P., Dawson, D.A., Horsburgh, G.J., Krupa, A.P. & Stevens, J.R. (2013a) Isolation and characterization of fourteen microsatellite loci from the endangered octocoral *Eunicella verrucosa* (Pallas 1766). *Conservation Genetics Resources* **5**, 825–829.

Holland, L.P., Dawson, D.a., Horsburgh, G.J. & Stevens, J.R. (2013b) Isolation and characterization of 11 microsatellite loci from the ubiquitous temperate octocoral *Alcyonium digitatum* (Linnaeus, 1758). *Conservation Genetics Resources* **5**, 767–770.

Holland, L.P., Jenkins, T.L. & Stevens, J.R. (2017) Contrasting patterns of population structure and gene flow facilitate exploration of connectivity in two widely distributed temperate octocorals. *Heredity* **119**, 35–48.

Huang, H. & Knowles, L.L. (2016) Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of rad sequences. *Systematic Biology* **65**, 357–365.

Hughes, A.L.C., Gyllencreutz, R., Lohne, Ø.S., Mangerud, J. & Svendsen, J.I. (2016) The last Eurasian ice sheets - a chronological database and time-slice reconstruction, DATED-1. *Boreas* **45**, 1–45.

Huserbraten, M.B.O., Moland, E., Knutsen, H., Olsen, E.M., André, C. & Stenseth, N.C. (2013) Conservation, spillover and gene flow within a network of northern European Marine Protected Areas. *PLoS ONE* **8**, e73388.

Jacob, S., Legrand, D., Chaine, A.S., Bonte, D., Schtickzelle, N., Huet, M. & Clobert, J. (2017) Gene flow favours local adaptation under habitat choice in ciliate microcosms. *Nature Ecology & Evolution* **1**, 1407–1410.

Jacobs, A., De Noia, M., Praebel, K., Kanstad-Hanssen, O., Paterno, M., Jackson, D., McGinnity, P., Sturm, A., Elmer, K.R. & Llewellyn, M.S. (2018) Genetic fingerprinting of salmon louse (*Lepeophtheirus salmonis*) populations in the North-East Atlantic using a random forest classification approach. *Scientific Reports* **8**, 1203.

Jahnke, M., Jonsson, P.R., Moksnes, P.O., Loo, L.O., Nilsson Jacobi, M. & Olsen, J.L. (2018) Seascape genetics and biophysical connectivity modelling support conservation of the seagrass *Zostera marina* in the Skagerrak-Kattegat region of the eastern North Sea. *Evolutionary Applications* **11**, 645–661.

Jakobsson, M. & Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806.

Janes, J.K., Miller, J.M., Dupuis, J.R., Malenfant, R.M., Gorrell, J.C., Cullingham, C.I. & Andrew, R.L. (2017) The $K = 2$ conundrum. *Molecular Ecology* **26**, 3594–3602.

Janko, K., Lecointre, G., Devries, A., Couloux, A., Cruaud, C. & Marshall, C. (2007) Did glacial advances during the Pleistocene influence differently the demographic histories of benthic and pelagic Antarctic shelf fishes? Inferences from intraspecific mitochondrial and nuclear DNA sequence diversity. *BMC Evolutionary Biology* **7**, 220.

Jenkins, T.L., Castilho, R. & Stevens, J.R. (2018a) Meta-analysis of northeast Atlantic marine taxa shows contrasting phylogeographic patterns following post-LGM expansions. *PeerJ* **6**, e5684.

Jenkins, T.L., Ellis, C.D. & Stevens, J.R. (2018b) SNP discovery in European lobster (*Homarus gammarus*) using RAD sequencing. *Conservation Genetics Resources* ***in press***, doi: 10.1007/s12686–018–1001–8.

Jenkins, T.L. & Stevens, J.R. (2018) Assessing connectivity between MPAs: selecting taxa and translating genetic data to inform policy. *Marine Policy* **94**, 165–173.

Jiao, W., Fu, X., Li, J., Li, L., Feng, L., Lv, J., Zhang, L., Wang, X., Li, Y., Hou, R., Zhang, L., Hu, X., Wang, S. & Bao, Z. (2014) Large-scale development of gene-associated single-nucleotide polymorphism markers for molluscan population genomic, comparative genomic, and genome-wide association studies. *DNA Research* **21**, 183–193.

Johannesson, K. & André, C. (2006) Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. *Molecular Ecology* **15**, 2013–2029.

Johnson, W.E., Onorato, D.P., Roelke, M.E., Land, E.D., Cunningham, M., Belden, R.C., McBride, R., Jansen, D., Lotz, M., Shindle, D., Howard, J., Wildt, D.E., Penfold, L.M., Hostetler, J.A., Oli, M.K. & O'Brien, S.J. (2010) Genetic restoration of the Florida panther. *Science* **329**, 1641–1645.

Jolly, M.T., Jollivet, D., Gentil, F., Thiébaut, E. & Viard, F. (2005) Sharp genetic break between Atlantic and English Channel populations of the polychaete *Pectinaria koreni*, along the north coast of France. *Heredity* **94**, 23–32.

Jolly, M.T., Viard, F., Gentil, F., Thiébaut, E. & Jollivet, D. (2006) Comparative phylogeography of two coastal polychaete tubeworms in the Northeast Atlantic supports shared history and vicariant events. *Molecular Ecology* **15**, 1841–1855.

Jombart, T. & Ahmed, I. (2011) adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071.

Jombart, T., Devillard, S. & Balloux, F. (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**, 94.

Jones, A.G., Small, C.M., Paczolt, K.A. & Ratterman, N.L. (2010) A practical guide to methods of parentage analysis. *Molecular Ecology Resources* **10**, 6–30.

Jones, G.P., Srinivasan, M. & Almany, G.R. (2007) Population connectivity and conservation of marine biodiversity. *Oceanography* **20**, 100–111.

Jørstad, K.E., Agnalt, A.L., Kristiansen, T.S. & Nøstvold, E. (2001) High survival and growth of European lobster juveniles (*Homarus gammarus*) reared communally on a natural-bottom substrate. *Marine and Freshwater Research* **52**, 1431–1438.

Jørstad, K.E., Faresteit, E., Kelly, E. & Triantaphyllids, C. (2005) Allozyme variation in European lobster (*Homarus gammarus*) throughout its distribution range. *New Zealand Journal of Marine and Freshwater Research* **39**, 515–526.

Jørstad, K.E., Kristiansen, T.S., Farestveit, E., Agnalt, A.L., Prodöhl, P.A., Hughes, M., Ferguson, A. & Taggert, J.B. (2009) Survival of laboratory-reared juvenile European lobster (*Homarus gammarus*) from three brood sources in southwestern Norway. *New Zealand Journal of Marine and Freshwater Research* **43**, 59–68.

Jost, L. (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology* **17**, 4015–4026.

Kaiser, M.J., Hormbrey, S., Booth, J.R., Hinz, H. & Hiddink, J.G. (2018) Recovery linked to life history of sessile epifauna following exclusion of towed mobile fishing gear. *Journal of Applied Ecology* **55**, 1060–1070.

Kalinkat, G., Cabral, J.S., Darwall, W., Ficetola, G.F., Fisher, J.L., Giling, D.P., Gosselin, M.P., Grossart, H.P., Jähnig, S.C., Jeschke, J.M., Knopf, K., Larsen, S., Onandia, G., Pätzig, M., Saul, W.C., Singer, G., Sperfeld, E. & Jarić, I. (2017) Flagship umbrella species needed for the conservation of overlooked aquatic biodiversity. *Conservation Biology* **31**, 481–485.

Kalinowski, S. (2002) How many alleles per locus should be used to estimate genetic distances? *Heredity* **88**, 62–65.

Kamvar, Z.N., Tabima, J.F. & Grünwald, N.J. (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2:**, e281.

Karl, S.A., Toonen, R.J., Grant, W.S. & Bowen, B.W. (2012) Common misconceptions in molecular ecology: echoes of the modern synthesis. *Molecular Ecology* **21**, 4171–4189.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649.

Keenan, K., Mcginnity, P., Cross, T.F., Crozier, W.W. & Prodohl, P.A. (2013) DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* **4**, 782–788.

Kelly, R.P. & Palumbi, S.R. (2010) Genetic structure among 50 species of the northeastern pacific rocky intertidal community. *PLoS ONE* **5**, e8594.

Kimura, M. & Weiss, G.H. (1964) The Stepping Stone Model of population structure and the decrease of genetic correlation with distance. *Genetics* **49**, 561–576.

Kleiven, A.R., Olsen, E.M. & Vølstad, J.H. (2012) Total catch of a red-listed marine species is an order of magnitude higher than official data. *PLoS ONE* **7**, e31216.

Kool, J.T., Moilanen, A. & Treml, E.A. (2013) Population connectivity: recent advances and new perspectives. *Landscape Ecology* **28**, 165–185.

Kough, A.S. & Paris, C.B. (2015) The influence of spawning periodicity on population connectivity. *Coral Reefs* **34**, 753–757.

Kough, A.S., Paris, C.B. & Butler IV, M.J. (2013) Larval connectivity and the international management of fisheries. *PLoS ONE* **8**, e64970.

Krakau, M., Jacobsen, S., Jensen, K.T. & Reise, K. (2012) The cockle *Cerastoderma edule* at northeast Atlantic shores: genetic signatures of glacial refugia. *Marine Biology* **159**, 221–230.

Krone, R. & Schröder, A. (2011) Wrecks as artificial lobster habitats in the German Bight. *Helgoland Marine Research* **65**, 11–16.

Laakkonen, H.M., Strelkov, P. & Väinölä, R. (2015) Molecular lineage diversity and inter-oceanic biogeographical history in *Hiatella* (Mollusca, Bivalvia). *Zoologica Scripta* **44**, 383–402.

Laikre, L., Schwartz, M.K., Waples, R.S. & Ryman, N. (2010) Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals. *Trends in Ecology and Evolution* **25**, 520–529.

Lambeck, K. & Chappell, J. (2001) Sea level change through the last glacial cycle. *Science* **292**, 679–686.

Lambeck, K., Esat, T.M. & Potter, E.K. (2002) Links between climate and sea levels for the past three million years. *Nature* **419**, 199–206.

Lárez, M., Palazón-Fernández, J. & Bolaños, C. (2000) The effect of salinity and temperature on the larval development of *Mithrax caribbaeus* Rathbun, 1920 (Brachyura: Majidae) reared in the laboratory. *Journal of Plankton Research* **22**, 1855–1868.

Larmuseau, M.H.D., Van Houdt, J.K.J., Guelinckx, J., Hellemans, B. & Volckaert, F.A.M. (2009) Distributional and demographic consequences of Pleistocene climate fluctuations for a marine demersal fish in the north-eastern Atlantic. *Journal of Biogeography* **36**, 1138–1151.

Lawson, D.J., van Dorp, L. & Falush, D. (2018) A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications* **9**, 3258.

Ledoux, J.B., Mokhtar-Jamai, K., Roby, C., Feral, J.P., Garrabou, J. & Aurelle, D. (2010) Genetic survey of shallow populations of the Mediterranean red coral [*Corallium rubrum* (Linnaeus, 1758)]: new insights into evolutionary processes shaping nuclear diversity and implications for conservation. *Molecular Ecology* **19**, 675–690.

Leenhardt, P., Cazalet, B., Salvat, B., Claudet, J. & Feral, F. (2013) The rise of large-scale marine protected areas: conservation or geopolitics? *Ocean and Coastal Management* **85**, 112–118.

Legendre, P., Fortin, M.J. & Borcard, D. (2015) Should the Mantel test be used in spatial analysis? *Methods in Ecology and Evolution* **6**, 1239–1247.

Lehnert, S.J., Dibacco, C., Wyngaarden, M.V., Jeffery, N.W., Lowen, J.B., Sylvester, E.V.A., Wringe, B.F., Stanley, R.R.E., Hamilton, L.C. & Bradbury, I.R. (2018) Fine-scale temperature-associated genetic structure between inshore and offshore populations of sea scallop (*Placopecten magellanicus*). *Heredity* ***in press***, https://doi.org/10.1038/s41437–018–0087–9.

Leis, J.M. (2006) Are larvae of demersal fishes plankton or nekton? **51**, 57–141.

Lessios, H. (2008) The Great American Schism: divergence of marine organisms after the rise of the Central American Isthmus. *Annual Review of Ecology, Evolution, and Systematics* **39**, 63–91.

Levin, L.A. (2006) Recent progress in understanding larval dispersal: new directions and digressions. *Integrative and Comparative Biology* **46**, 282–297.

Li, Y., Wang, W., Liu, X., Luo, W., Zhang, J. & Gul, Y. (2011) DNA extraction from crayfish exoskeleton. *Indian Journal of Experimental Biology* **49**, 953–957.

Librado, P. & Rozas, J. (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452.

Lieberknecht, L.M. & Jones, P.J.S. (2016) From stormy seas to the doldrums: the challenges of navigating towards an ecologically coherent marine protected area network through England's Marine Conservation Zone process. *Marine Policy* **71**, 275–284.

Lien, V.S., Gusdal, Y. & Vikebø, F.B. (2014) Along-shelf hydrographic anomalies in the Nordic Seas (1960-2011): locally generated or advective signals? *Ocean Dynamics* **64**, 1047–1059.

Liggins, L., Treml, E.A., Possingham, H.P. & Riginos, C. (2016) Seascape features, rather than dispersal traits, predict spatial genetic patterns in co-distributed reef fishes. *Journal of Biogeography* **43**, 256–267.

Linnane, A., Ball, B., Mercer, J.P., Browne, R., Meeren, G.V.D., Ringvold, H., Bannister, C., Mazzoni, D. & Munday, B. (2001) Searching for the early benthic phase (EBP) of the European lobster: a trans-European study of cobble fauna. *Hydrobiologia* **465**, 63–72.

Lorenzen, K., Beveridge, M.C. & Mangel, M. (2012) Cultured fish: integrative biology and management of domestication and interactions with wild fish. *Biological Reviews* **87**, 639–660.

Lotterhos, K.E. & Whitlock, M.C. (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology* **24**, 1031–1046.

Lowe, W.H. & Allendorf, F.W. (2010) What can genetics tell us about population connectivity? *Molecular Ecology* **19**, 3038–3051.

Lowry, E., Rollinson, E.J., Laybourn, A.J., Scott, T.E., Aiello-Lammens, M.E., Gray, S.M., Mickley, J. & Gurevitch, J. (2013) Biological invasions: a field synopsis, systematic review, and database of the literature **3**, 182–196.

Lubchenco, J. (2016) Ecology: the sea-otter whisperer. *Nature* **533**, 318–319.

Luu, K., Bazin, E. & Blum, M.G.B. (2017) pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources* **17**, 67–77.

Maggs, C., Castilho, R., Foltz, D., Henzler, C., Jolly, M., Kelly, J., Olsen, J., Perez, K., Stam, W., Vainola, R., Viard, F. & Wares, J. (2008) Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. *Ecology* **89**, 108–122.

Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J. & Turner, D.J. (2010) Target-enrichment strategies for next-generation sequencing.

Manel, S., Gaggiotti, O.E. & Waples, R.S. (2005) Assignment methods: matching biological questions with appropriate techniques. *Trends in Ecology and Evolution* **20**, 136–142.

Marandel, F., Lorance, P., Berthelé, O., Trenkel, V.M., Waples, R.S. & Lamy, J.B. (2018) Estimating effective population size of large marine populations, is it feasible? *Fish and Fisheries* **in press**, https://doi.org/10.1111/faf.12338.

Mardis, E.R. (2011) A decade's perspective on DNA sequencing technology. *Nature* **470**, 198–203.

Marko, P.B., Hoffman, J.M., Emme, S.A., McGovern, T.M., Keever, C.C. & Nicole Cox, L. (2010) The "Expansion-Contraction" model of Pleistocene biogeography: rocky shores suffer a sea change? *Molecular Ecology* **19**, 146–169.

Maroso, F., Casanova, A., do Prado, F.D., Bouza, C., Pardo, B.G., Blanco, A., Hermida, M., Fernández, C., Vera, M. & Martínez, P. (2018) Species identification of two closely exploited flatfish, turbot (*Scophthalmus maximus*) and brill (*Scophthalmus rhombus*), using a ddRADseq genomic approach. *Aquatic Conservation: Marine and Freshwater Ecosystems* **28**, 1253–1260.

Marti-Puig, P., Costantini, F., Rugiu, L., Ponti, M. & Abbiati, M. (2013) Patterns of genetic connectivity in invertebrates of temperate MPA networks. *Advances in Oceanography and Limnology* **4**, 138–149.

Martinsohn, J.T. & Ogden, R. (2009) FishPopTrace-Developing SNP-based population genetic assignment methods to investigate illegal fishing. *Forensic Science International: Genetics Supplement Series* **2**, 294–296.

Masmoudi, M.B., Chaoui, L., Topçu, N.E., Hammami, P., Kara, M.H. & Aurelle, D. (2016) Contrasted levels of genetic diversity in a benthic Mediterranean octocoral: consequences of different demographic histories? *Ecology and Evolution* **6**, 8665–8678.

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T.H., Piñero, D. & Emerson, B.C. (2014) Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference. *Molecular Ecology Resources* pp. 1–14.

McCormick, M.I. & Moloney, B.W. (1995) Influence of water temperature during the larval stage on size, age and body condition of a tropical reef fish at settlement. *Marine Ecology Progress Series* **118**, 59–68.

Meek, M.H., Baerwald, M.R., Stephens, M.R., Goodbla, A., Miller, M.R., Tomalty, K.M. & May, B. (2016) Sequencing improves our ability to study threatened migratory species: genetic population assignment in California's Central Valley Chinook salmon. *Ecology and Evolution* **6**, 7706–7716.

Meirmans, P.G. (2015) Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology* **24**, 3223–3231.

Ménot, G., Bard, E., Rostek, F., Weijers, J.W.H., Hopmans, E.C., Schouten, S. & Sinninghe Damsté, J.S. (2006) Early reactivation of European rivers during the last deglaciation. *Science* **313**, 1623–1625.

Mertens, L.E.A., Treml, E.A. & von der Heyden, S. (2018) Genetic and Biophysical Models Help Define Marine Conservation Focus Areas. *Frontiers in Marine Science* ***in press***, https://doi.org/10.3389/fmars.2018.00268.

Metaxas, A. & Saunders, M. (2009) Quantifying the "Bio-" components in biophysical models of larval transport in marine benthic invertebrates: advances and pitfalls. *Biological Bulletin* **216**, 257–272.

Metzker, M.L. (2010) Sequencing technologies - the next generation. *Nature Reviews Genetics* **11**, 31–46.

Miller, M., Dunham, J., Amores, A., Cresko, W. & Johnson, E. (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD). *Genome Research* **17**, 240–248.

Mokhtar-Jamai, K., Pascual, M., Ledoux, J.B., Coma, R., Feral, J.P., Garrabou, J. & Aurelle, D. (2011) From global to local genetic structuring in the red gorgonian *Paramuricea clavata*: the interplay between oceanographic conditions and limited larval dispersal. *Molecular Ecology* **20**, 3291–3305.

Moland, E., Olsen, E.M., Knutsen, H., Garrigou, P., Espeland, S.H., Kleiven, A.R., André, C., Knutsen, J.A., Kleiven, R., B, P.R.S. & Andre, C. (2013) Lobster and cod benefit from small-scale northern marine protected areas: inference from an empirical before-after control-impact study. *Proceedings of the Royal Society B* **280**, 20122679.

Morin, P.A., Martien, K.K. & Taylor, B.L. (2009) Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources* **9**, 66–73.

Moritz, C. (1994) Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology and Evolution* **9**, 373–375.

Moritz, C. (1999) Conservation units and traslocations: stategies for conserving evolutionary processes. *Hereditas* **130**, 217–228.

Morjan, C.L. & Rieseberg, L.H. (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology* **13**, 1341–1356.

Mullins, R.B., McKeown, N.J., Sauer, W.H.H. & Shaw, P.W. (2018) Genomic analysis reveals multiple mismatches between biological and management units in yellowfin tuna (*Thunnus albacares*). *ICES Journal of Marine Science* ***in press***, https://doi.org/10.1093/icesjms/fsy102.

Munro, L. (2004) Determining the reproductive cycle of *Eunicella verrucosa*. RR Report 07/2004 ETR 12. Tech. rep.

Nanninga, G.B. & Manica, A. (2018) Larval swimming capacities affect genetic differentiation and range size in demersal marine fishes. *Marine Ecology Progress Series* **589**, 1–12.

Nei, M. & Chesser, R. (1983) Estimation of fixation indices and gene diversities. *Annals of Human Genetics* **47**, 253–257.

Neiva, J., Assis, J., Fernandes, F., Pearson, G.A. & Serrão, E.A. (2014) Species distribution models and mitochondrial DNA phylogeography suggest an extensive biogeographical shift in the high-intertidal seaweed *Pelvetia canaliculata*. *Journal of Biogeography* **41**, 1137–1148.

Neiva, J., Pearson, G.A., Valero, M. & Serrão, E.A. (2012) Drifting fronds and drifting alleles: range dynamics, local dispersal and habitat isolation shape the population structure of the estuarine seaweed *Fucus ceranoides*. *Journal of Biogeography* **39**, 1167–1178.

Ni, G., Li, Q., Kong, L. & Yu, H. (2014) Comparative phylogeography in marginal seas of the northwestern Pacific. *Molecular Ecology* **23**, 534–548.

Ni, G., Li, Q., Kong, L. & Zheng, X. (2012) Phylogeography of the bivalve *Tegillarca granosa* in coastal China: Implications for management and conservation. *Marine Ecology Progress Series* **452**, 119–130.

Nielsen, E.E., Cariani, A., Aoidh, E.M., Maes, G.E., Milano, I., Ogden, R., Taylor, M., Hemmer-Hansen, J., Babbucci, M., Bargelloni, L., Bekkevold, D., Diopere, E., Grenfell, L., Helyar, S., Limborg, M.T., Martinsohn, J.T., McEwing, R., Panitz, F., Patarnello, T., Tinti, F., Van Houdt, J.K., a.M. Volckaert, F., Waples, R.S., Albin, J.E., Vieites Baptista, J.M., Barmintsev, V., Bautista, J.M., Bendixen, C., Bergé, J.P., Blohm, D., Cardazzo, B., Diez, A., Espiñeira, M., Geffen, A.J., Gonzalez, E., González-Lavín, N., Guarniero, I., Jeráme, M., Kochzius, M., Krey, G., Mouchel, O., Negrisolo, E., Piccinetti, C., Puyet, A., Rastorguev, S., Smith, J.P., Trentini, M., Verrez-Bagnis, V., Volkov, A., Zanzi, A. & Carvalho, G.R. (2012) Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature Communications* p. 3:851.

Nielsen, R. (2005) Molecular signatures of natural selection. *Annual Review of Genetics* **39**, 197–218.

Nunziata, S.O. & Weisrock, D.W. (2018) Estimation of contemporary effective population size and population declines using RAD sequence data. *Heredity* **120**, 196–207.

O'Connor, M.I., Bruno, J.F., Gaines, S.D., Halpern, B.S., Lester, S.E., Kinlan, B.P. & Weiss, J.M. (2007) Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proceedings of the National Academy of Sciences* **104**, 1266–1271.

O'Dea, A., Lessios, H.A., Coates, A.G., Eytan, R.I., Restrepo-Moreno, S.A., Cione, A.L., Collins, L.S., de Queiroz, A., Farris, D.W., Norris, R.D., Stallard, R.F., Woodburne, M.O., Aguilera, O., Aubry, M.P., Berggren, W.A., Budd, A.F., Cozzuol, M.A., Coppard, S.E., Duque-Caro, H., Finnegan, S., Gasparini, G.M., Grossman, E.L., Johnson, K.G., Keigwin, L.D., Knowlton, N., Leigh, E.G., Leonard-Pingel, J.S., Marko, P.B., Pyenson, N.D., Rachello-Dolmen, P.G., Soibelzon, E., Soibelzon, L., Todd, J.A., Vermeij, G.J. & Jackson, J.B.C. (2016) Formation of the Isthmus of Panama. *Science Advances* **2**, 1–11.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E. & Wagner, H. (2018) *vegan: Community Ecology Package*. R package version 2.5-2.

O'Leary, B.C., Ban, N.C., Fernandez, M., Friedlander, A.M., García-Borboroglu, P., Golbuu, Y., Guidetti, P., Harris, J.M., Hawkins, J.P., Langlois, T., McCauley, D.J., Pikitch, E.K., Richmond, R.H. & Roberts, C.M. (2018) Addressing criticisms of large-scale Marine Protected Areas. *BioScience* **68**, 359–370.

Olsen, J.L., Zechman, F.W., Hoarau, G., Coyer, J.A., Stam, W.T., Valero, M. & Åberg, P. (2010) The phylogeographic architecture of the fucoid seaweed *Ascophyllum nodosum*: an intertidal 'marine tree' and survivor of more than one glacial-interglacial cycle. *Journal of Biogeography* **37**, 842–856.

Oresland, V. & Ulmestrand, M. (2013) European lobster subpopulations from limited adult movements and larval retention. *ICES Journal of Marine Science* **70**, 532–539.

OSPAR Commission (2007) Background document to support the assessment of whether the OSPAR Network of Marine Protected Areas is ecologically coherent pp. ISBN: 978–1–905859–59–7.

OSPAR Commission (2010) Results of the e-consultation on the draft quality status report 2010 pp. ISBN 978–1–907390–41–8.

OSPAR Commission (2013) An assessment of the ecological coherence of the OSPAR Network of Marine Protected Areas in 2012, pp. ISBN: 978–1–909159–52–5.

OSPAR Commission (2017) 2016 Status Report on the OSPAR Network of Marine Protected Areas pp. ISBN: 978–1–911458–33–3.

Ovenden, J.R. (2013) Crinkles in connectivity: combining genetics and other types of biological data to estimate movement and interbreeding between populations. *Marine and Freshwater Research* **64**, 201–207.

Ozsolak, F. & Milos, P.M. (2011) RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics* **12**, 87–89.

Palero, F., Abelló, P., Macpherson, E., Beaumont, M. & Pascual, M. (2011) Effect of oceanographic barriers and overfishing on the population genetic structure of the European spiny lobster (*Palinurus elephas*). *Biological Journal of the Linnean Society* **104**, 407–418.

Palero, F., Abelló, P., Macpherson, E., Gristina, M. & Pascual, M. (2008) Phylogeography of the European spiny lobster (*Palinurus elephas*): influence of current oceanographical features and historical processes. *Molecular Phylogenetics and Evolution* **48**, 708–717.

Palsbøll, P.J., Bérubé, M. & Allendorf, F.W. (2007) Identification of management units using population genetic data. *Trends in Ecology and Evolution* **22**, 11–16.

Palumbi, S.R. (2003) Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications* **13**, 146–158.

Pante, E., Abdelkrim, J., Viricel, A., Gey, D., France, S.C., Boisselier, M.C. & Samadi, S. (2015a) Use of RAD sequencing for delimiting species. *Heredity* **114**, 450–459.

Pante, E., Puillandre, N., Viricel, A., Arnaud-Haond, S., Aurelle, D., Castelin, M., Chenuil, A., Destombe, C., Forcioli, D., Valero, M., Viard, F. & Samadi, S. (2015b) Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Molecular Ecology* **24**, 525–544.

Pante, E. & Simon-Bouhet, B. (2013) marmap: a Package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE* **8**, e73051.

Paradis, E. (2010) Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**, 419–420.

Paris, J.R., Stevens, J.R. & Catchen, J.M. (2017) Lost in parameter space: a road map for stacks. *Methods in Ecology and Evolution* **8**, 1360–1373.

Pascual, M., Rives, B., Schunter, C. & Macpherson, E. (2017) Impact of life history traits on gene flow: a multispecies systematic review across oceanographic barriers in the Mediterranean Sea. *PLoS ONE* **12**, e0176419.

Patarnello, T., Volckaert, F.A.M.J. & Castilho, R. (2007) Pillars of Hercules: is the Atlantic-Mediterranean transition a phylogeographical break? *Molecular Ecology* **16**, 4426–4444.

Pechenik, J.A. (1990) Delayed metamorphosis by larvae of benthic marine invertebrates: Does it occur? Is there a price to pay? *Ophelia* **32**, 64–94.

Pechenik, J.A. (2006) Larval experience and latent effects - Metamorphosis is not a new beginning. *Integrative and Comparative Biology* **46**, 323–333.

Peijnenburg, K.T.C.A., Fauvelot, C., Breeuwer, J.A.J. & Menken, S.B.J. (2006) Spatial and temporal genetic structure of the planktonic *Sagitta setosa* (Chaetognatha) in European seas as revealed by mitochondrial and nuclear DNA markers. *Molecular Ecology* **15**, 3319–38.

Pelc, R.A., Warner, R.R. & Gaines, S.D. (2009) Geographical patterns of genetic structure in marine species with contrasting life histories. *Journal of Biogeography* **36**, 1881–1890.

Perez, M.F., Franco, F.F., Bombonato, J.R., Bonatelli, I.A.S., Khan, G., Romeiro-Brito, M., Fegies, A.C., Ribeiro, P.M., Silva, G.A.R. & Moraes, E.M. (2018) Assessing population structure in the face of isolation by distance: are we neglecting the problem? *Diversity and Distributions* **24**, 1883–1889.

Pérez-Losada, M., Guerra, A., Carvalho, G.R., Sanjuan, A. & Shaw, P.W. (2002) Extensive population subdivision of the cuttlefish *Sepia officinalis* (Mollusca: Cephalopoda) around the Iberian Peninsula indicated by microsatellite DNA variation. *Heredity* **89**, 417–424.

Pérez-Portela, R., Bumford, A., Coffman, B., Wedelich, S., Davenport, M., Fogg, A., Swenarton, M.K., Coleman, F., Johnston, M.A., Crawford, D.L. & Oleksiak, M.F. (2018) Genetic homogeneity of the invasive lionfish across the Northwestern Atlantic and the Gulf of Mexico based on single nucleotide polymorphisms. *Scientific Reports* **8**, 5062.

Pérez-Portela, R., Turon, X. & Bishop, J. (2012) Bottlenecks and loss of genetic diversity: spatio-temporal patterns of genetic structure in an ascidian recently introduced in Europe. *Marine Ecology Progress Series* **451**, 93–105.

Pikesley, S.K., Godley, B.J., Latham, H., Richardson, P.B., Robson, L.M., Solandt, J.L., Trundle, C., Wood, C. & Witt, M.J. (2016) Pink sea fans (*Eunicella verrucosa*) as indicators of the spatial efficacy of Marine Protected Areas in southwest UK coastal waters. *Marine Policy* **64**, 38–45.

Pillai, S., Gopalan, V. & Lam, A.K.Y. (2017) Review of sequencing platforms and their applications in phaeochromocytoma and paragangliomas. *Critical Reviews in Oncology/Hematology* **116**, 58–67.

Pita, A., Presa, P. & Perez, M. (2010) Gene flow, multilocus assignment and genetic structuring of the European hake (*Merluccius merluccius*). *Thalassas* **26**, 129–133.

Pivotto, I.D., Nerini, D., Masmoudi, M., Kara, H., Chaoui, L. & Aurelle, D. (2015) Highly contrasted responses of Mediterranean octocorals to climate change along a depth gradient. *Royal Society Open Science* **2**, 140493.

Podbielski, I., Bock, C., Lenz, M. & Melzner, F. (2016) Using the critical salinity (Scrit) concept to predict invasion potential of the anemone Diadumene lineata in the Baltic Sea. *Marine Biology* **163**, 227.

Pop, M. & Salzberg, S.L. (2008) Bioinformatics challenges of new sequencing technology. *Trends in Genetics* **24**, 142–149.

Pratlong, M., Rancurel, C., Pontarotti, P. & Aurelle, D. (2016) Monophyly of Anthozoa (Cnidaria): why do nuclear and mitochondrial phylogenies disagree? *Zoologica Scripta* **46**, 363–371.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Provan, J. & Bennett, K.D. (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution* **23**, 564–571.

Provan, J., Wattier, R.A. & Maggs, C.A. (2005) Phylogeographic analysis of the red seaweed *Palmaria palmata* reveals a Pleistocene marine glacial refugium in the English Channel. *Molecular Ecology* **14**, 793–803.

Puckett, B.J., Eggleston, D.B., Kerr, P.C. & Luettich, R.A. (2014) Larval dispersal and population connectivity among a network of marine reserves. *Fisheries Oceanography* **23**, 342–361.

Puebla, O. (2018) Another useful property of mtDNA: editorial comment on the highlighted article by Lou et al. (2018). *Marine Biology* **165**, 125.

Puechmaille, S. (2016) The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: sub-sampling and new estimators alleviate the problem. *Molecular Ecology Resources* **16**, 608–627.

Pysek, P. & Richardson, D.M. (2010) Invasive species, environmental change, and health. *Annual Review of Environment and Resources* **35**, 25–55.

Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P. & Gu, Y. (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* **13**, 341.

Quinn, B.K., Rochette, R., Ouellet, P. & Sainte-Marie, B. (2013) Effect of temperature on development rate of larvae from cold-water American lobster (*Homarus americanus*). *Journal of Crustacean Biology* **33**, 527–536.

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/ .

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**, 901–904.

Ramos-Onsins, S.E. & Rozas, J. (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19**, 2092–2100.

Reiss, H., Hoarau, G., Dickey-Collas, M. & Wolff, W.J. (2009) Genetic population structure of marine fish: mismatch between biological and fisheries management units. *Fish and Fisheries* **10**, 361–395.

Reitzel, A.M., Herrera, S., Layden, M.J., Martindale, M.Q. & Shank, T.M. (2013) Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular Ecology* **22**, 2953–2970.

Remerie, T., Vierstraete, A., Weekers, P.H.H., Vanfleteren, J.R. & Vanreusel, A. (2009) Phylogeography of an estuarine mysid, *Neomysis integer* (Crustacea, Mysida), along the north-east Atlantic coasts. *Journal of Biogeography* **36**, 39–54.

Reuschel, S., Cuesta, J.A. & Schubart, C.D. (2010) Marine biogeographic boundaries and human introduction along the European coast revealed by phylogeography of the prawn *Palaemon elegans*. *Molecular Phylogenetics and Evolution* **55**, 765–775.

Riginos, C., Crandall, E.D., Liggins, L., Bongaerts, P. & Treml, E.A. (2016) Navigating the currents of seascape genomics: how spatial analyses can augment population genomic studies. *Current Zoology* **62**, 581–601.

Riginos, C., Douglas, K.E., Jin, Y., Shanahan, D.F. & Treml, E.A. (2011) Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography* **34**, 566–575.

Riquet, F., Liautard-Haag, C., Woodall, L., Bouza, C., Louisy, P., Hamer, B., Otero-Ferrer, F., Aublanc, P., Beduneau, V., Briard, O., Ayari, T.E., Hochscheid, S., Belkhir, K., Arnaud-Haond, S., Gagnaire, P.A. & Bierne, N. (2017) Parallel use of a shared genomic island of speciation in clinal and mosaic hybrid zones between cryptic seahorse lineages. *bioRxiv* **161786**, ver. 4.

Roach, M., Cohen, M., Forster, R., Revill, A.S. & Johnson, M. (2018) The effects of temporary exclusion of activity due to wind farm construction on a lobster (*Homarus gammarus*) fishery suggests a potential management approach. *ICES Journal of Marine Science* **75**, 1416–1426.

Robalo, J.I., Castilho, R., Francisco, S.M., Almada, F., Knutsen, H., Jorde, P.E., Pereira, A.M. & Almada, V.C. (2012) Northern refugia and recent expansion in the North Sea: the case of the wrasse *Symphodus melops* (Linnaeus, 1758). *Ecology and Evolution* **2**, 153–164.

Roberts, C.M., Hawkins, J.P., Fletcher, J., Hands, S., Raab, K. & Ward, S. (2010) Guidance on the size and spacing of Marine Protected Areas in England. Tech. rep., Natural England, Commissioned Report NECR037, Natural England, Commissioned Report NECR037.

Rohling, E.J., Fenton, M., Jorissen, F.J., Bertrand, P., Ganssen, G. & Caulet, J.P. (1998) Magnitudes of sea-level lowstands of the past 500,000 years. *Nature* **394**, 162–165.

Roman, J. & Palumbi, S.R. (2004) A global invader at home: population structure of the green crab, *Carcinus maenas*, in Europe. *Molecular Ecology* **13**, 2891–2898.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N., Loire, E., Lourenco, J.M., Nabholz, B., Roux, C., Tsagkogeorga, G., Weber, A.A.T., Weinert, L.A., Belkhir, K., Bierne, N., Glémin, S. & Galtier, N. (2014) Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* **515**, 261–263.

Rowe, G., Sweet, M. & Beebee, T. (2017) *An Introduction to Molecular Ecology*. Oxford University Press, Oxford, United Kingdom.

Rowe, H.C., Renaut, S. & Guggisberg, A. (2011) RAD in the realm of next-generation sequencing technologies. *Molecular Ecology* **20**, 3499–3502.

Runge, J., Runge, M. & Nichols, J. (2006) The role of local populations within a landscape context: defining and classifying sources and sinks. *The American Naturalist* **167**, 925–938.

Russello, M.A., Waterhouse, M.D., Etter, P.D. & Johnson, E.A. (2015) From promise to practice: pairing non-invasive sampling with genomics in conservation. *PeerJ* **3**, e1106.

Ryder, O.A. (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology & Evolution* **1**, 9–10.

Sá-Pinto, A., Branco, M.S., Alexandrino, P.B., Fontaine, M.C. & Baird, S.J.E. (2012) Barriers to gene flow in the marine environment: insights from two common intertidal limpet species of the Atlantic and Mediterranean. *PloS one* **7**, e50330.

Saenz-Agudelo, P., Jones, G.P., Thorrold, S.R. & Planes, S. (2009) Estimating connectivity in marine populations: an empirical evaluation of assignment tests and parentage analysis under different gene flow scenarios. *Molecular Ecology* **18**, 1765–1776.

Sandoval-Castillo, J., Robinson, N.A., Hart, A.M., Strain, L.W. & Beheregaray, L.B. (2018) Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (*Haliotis laevigata*), along a longitudinal environmental gradient. *Molecular Ecology* **27**, 1603–1620.

Sanford, E. & Kelly, M.W. (2011) Local adaptation in marine invertebrates. *Annual Review of Marine Science* **3**, 509–535.

Sartoretto, S. & Francour, P. (2012) Bathymetric distribution and growth rates of *Eunicella verrucosa* (Cnidaria: Gorgoniidae) populations along the Marseilles coast (France). *Scientia Marina* **76**, 349–355.

Saura, S., Bodin, Ö. & Fortin, M.J. (2014) EDITOR'S CHOICE: Stepping stones are crucial for species' long-distance dispersal and range expansion through habitat networks. *Journal of Applied Ecology* **51**, 171–182.

Schlötterer, C. (2004) The evolution of molecular markers - just a matter of fashion? *Nature Reviews Genetics* **5**, 63–69.

Schmalenbach, I. & Buchholz, F. (2010) Vertical positioning and swimming performance of lobster larvae (*Homarus gammarus*) in an artificial water column at Helgoland, North Sea. *Marine Biology Research* **6**, 89–99.

Schmalenbach, I. & Franke, H.D. (2010) Potential impact of climate warming on the recruitment of an economically and ecologically important species, the European lobster (*Homarus gammarus*) at Helgoland, North Sea. *Marine Biology* **157**, 1127–1135.

Schmalenbach, I., Mehrtens, F., Janke, M. & Buchholz, F. (2011) A mark-recapture study of hatchery-reared juvenile European lobsters, *Homarus gammarus*, released at the rocky island of Helgoland (German Bight, North Sea) from 2000 to 2009. *Fisheries Research* **108**, 22–30.

Seeb, J.E., Carvalho, G., Hauser, L., Naish, K., Roberts, S. & Seeb, L.W. (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources* **11**, 1–8.

Selkoe, K.A., D'Aloia, C.C., Crandall, E.D., Iacchei, M., Liggins, L., Puritz, J.B., Von Der Heyden, S. & Toonen, R.J. (2016) A decade of seascape genetics: contributions to basic and applied marine connectivity. *Marine Ecology Progress Series* **554**, 1–19.

Selkoe, K.A. & Toonen, R.J. (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series* **436**, 291–305.

Sellis, D., Callahan, B.J., Petrov, D.A. & Messer, P.W. (2011) Heterozygote advantage as a natural consequence of adaptation in diploids. *Proceedings of the National Academy of Sciences* **108**, 20666–20671.

Shafer, A.B.A., Cullingham, C.I., Côté, S.D. & Coltman, D.W. (2010) Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology* **19**, 4589–4621.

Shafer, A.B.A., Wolf, J.B.W., Alves, P.C., Bergstrom, L., Bruford, M.W., Brannstrom, I., Colling, G., Dalen, L., De Meester, L., Ekblom, R., Fawcett, K.D., Fior, S., Hajibabaei, M., Hill, J.A., Hoezel, A.R., Hoglund, J., Jensen, E.L., Krause, J., Kristensen, T.N., Krutzen, M., McKay, J.K., Norman, A.J., Ogden, R., Osterling, E.M., Ouborg, N.J., Piccolo, J., Popovic, D., Primmer, C.R., Reed, F.A., Roumet, M., Salmona, J., Schenekar, T., Schwartz, M.K., Segelbacher, G., Senn, H., Thaulow, J., Valtonen, M., Veale, A., Vergeer, P., Vijay, N., Vila, C., Weissensteiner, M., Wennerstrom, L., Wheat, C.W. & Zielinski, P. (2015) Genomics and the challenging translation into conservation practice. *Trends in Ecology and Evolution* **30**, 78–87.

Shanks, A.L. (2009) Pelagic larval duration and dispersal distance revisited. *Biological Bulletin* **216**, 373–385.

Shen, H., Braband, A. & Scholtz, G. (2015) The complete mitogenomes of lobsters and crayfish (Crustacea: Decapoda: Astacidea) reveal surprising differences in closely related taxa and convergences to Priapulida. *Journal of Zoological Systematics and Evolutionary Research* **53**, 273–281.

Sheppard, C.R.C., Ateweberhan, M., Bowen, B.W., Carr, P., Chen, C.A., Clubbe, C., Craig, M.T., Ebinghaus, R., Eble, J., Fitzsimmons, N., Gaither, M.R., Gan, C.H., Gollock, M., Guzman, N., Graham, N.A.J., Harris, A., Jones, R., Keshavmurthy, S., Koldewey, H., Lundin, C.G., Mortimer, J.A., Obura, D., Pfeiffer, M., Price, A.R.G., Purkis, S., Raines, P., Readman, J.W., Riegl, B., Rogers, A., Schleyer, M., Seaward, M.R.D., Sheppard, A.L.S., Tamelander, J., Turner, J.R., Visram, S., Vogler, C., Vogt, S., Wolschke, H., Yang, J.M.C., Yang, S.Y. & Yesson, C. (2012) Reefs and islands of the Chagos Archipelago, Indian Ocean: why it is the world's largest no-take marine protected area. *Aquatic Conservation: Marine and Freshwater Ecosystems* **22**, 232–261.

Siegel, D.A., Kinlan, B.P., Gaylord, B. & Gaines, S.D. (2003) Lagrangian descriptions of marine larval dispersion. *Marine Ecology Progress Series* **260**, 83–96.

Simberloff, D. (1998) Flagships, umbrellas, and keystones: is single-species management passé in the landscape era? *Biological Conservation* **83**, 247–257.

Skerritt, D., Robertson, P., Mill, A., Polunin, N. & Fitzsimmons, C. (2015) Fine-scale movement, activity patterns and home-ranges of European lobster *Homarus gammarus*. *Marine Ecology Progress Series* **536**, 203–219.

Slatkin, M. & Voelm, L. (1991) F_{ST} in a Hierarchial Island Model. *Genetics* **127**, 627–629.

Sørdalen, T.K., Halvorsen, K.T., Harrison, H.B., Ellis, C., Vøllestad, L.A., Knutsen, H., Moland, E. & Olsen, E.M. (2018) Harvesting changes mating behavior in European lobster. *Evolutionary Applications* **11**, 963–977.

Sotelo, G., Morán, P., Fernández, L. & Posada, D. (2008) Genetic variation of the spiny spider crab *Maja brachydactyla* in the northeastern Atlantic. *Marine Ecology Progress Series* **362**, 211–223.

Spanier, E., Lavalli, K.L., Goldstein, J.S., Groeneveld, J.C., Jordaan, G.L., Mallol, S., Jones, C.M., Phillips, B.F., Bianchini, M.L., Kibler, R.D., Díaz, D., Mallol, S., Goñi, R., van Der Meeren, G.I., Agnalt, A.L., Behringer, D.C., Keegan, W.F. & Jeffs, A. (2015) Populations from prehistory to the modern era. *ICES Journal of Marine Science* **72**, i7–i21.

Sponaugle, S., Lee, T., Kourafalou, V. & Pinkard, D. (2005) Florida Current frontal eddies and the settlement of coral reef fishes. *Limnology and Oceanography* **50**, 1033–1048.

Stiebens, V.A., Merino, S.E., Roder, C., Chain, F.J., Lee, P.L. & Eizaguirre, C. (2013) Living on the edge: how philopatry maintains adaptive potential. *Proceedings of the Royal Society B: Biological Sciences* **280**, 20130305.

Sunnucks, P. (2000) Efficient genetic markers for population biology. *Trends in Ecology and Evolution* **15**, 199–203.

Taberlet, P., Fumagalli, L., Wust-Saucy, A. & Cossons, J. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7**, 453–464.

Taboada, S. & Pérez-Portela, R. (2016) Contrasted phylogeographic patterns on mitochondrial DNA of shallow and deep brittle stars across the Atlantic-Mediterranean area. *Scientific Reports* **6**, 32425.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.

Taylor, H.R., Dussex, N. & van Heezik, Y. (2017) Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners. *Global Ecology and Conservation* **10**, 231–242.

Ten-Kate, K. (2002) Science and the convention on biological diversity. *Science* **295**, 2371–2372.

Thomas, L. & Bell, J.J. (2013) Testing the consistency of connectivity patterns for a widely dispersing marine species. *Heredity* **111**, 345–54.

Thorrold, S., Zacherl, D. & Levin, L. (2007) Population connectivity and larval dispersal using geochemical signatures in calcified structures. *Oceanography* **20**, 80–89.

Tintore, J., La Violette, P.E., Blade, I. & Cruzado, A. (1988) A study of an intense density front in the Eastern Alboran Sea: The Almeria-Oran Front. *Journal of Physical Oceanography* **18**, 1384–1397.

Todd, E.V., Black, M.A. & Gemmell, N.J. (2016) The power and promise of RNA-seq in ecology and evolution. *Molecular Ecology* **25**, 1224–1241.

Toonen, R.J., Andrews, K.R., Baums, I.B., Bird, C.E., Concepcion, G.T., Daly-Engel, T.S., Eble, J.A., Faucci, A., Gaither, M.R., Iacchei, M., Puritz, J.B., Schultz, J.K., Skillings, D.J., Timmers, M.a. & Bowen, B.W. (2011) Defining boundaries for ecosystem-based management: a multispecies case study of marine connectivity across the Hawaiian Archipelago. *Journal of Marine Biology* **2011**, 460173.

Toonen, R.J., Wilhelm, T.A., Maxwell, S.M., Wagner, D., Bowen, B.W., Sheppard, C.R., Taei, S.M., Teroroko, T., Moffitt, R., Gaymer, C.F., Morgan, L., Lewis, N., Sheppard, A.L., Parks, J. & Friedlander, A.M. (2013) One size does not fit all: the emerging frontier in large-scale marine conservation. *Marine Pollution Bulletin* **77**, 7–10.

Treml, E.A., Ford, J.R., Black, K.P. & Swearer, S.E. (2015) Identifying the key biophysical drivers, connectivity outcomes, and metapopulation consequences of larval dispersal in the sea. *Movement Ecology* **3**, 17.

Treml, E.A., Roberts, J.J., Chao, Y.Y., Halpin, P.N., Possingham, H.P. & Riginos, C. (2012) Reproductive output and duration of the pelagic larval stage determine seascape-wide connectivity of marine populations. *Integrative and Comparative Biology* **52**, 525–537.

Triantafyllidis, A., Apostolidis, A.P., Katsares, V., Kelly, E., Mercer, J., Hughes, M., Jorstad, K.E., Tsolou, A., Hynes, R. & Triantaphyllidis, C. (2005) Mitochondrial DNA variation in the European lobster (*Homarus gammarus*) throughout the range. *Marine Biology* **146**, 223–235.

Tripp, E.A., Tsai, Y.H.E., Zhuang, Y. & Dexter, K.G. (2017) RADseq dataset with 90% missing data fully resolves recent radiation of *Petalidium* (Acanthaceae) in the ultra-arid deserts of Namibia. *Ecology and Evolution* **7**, 7920–7936.

Tully, O., Roantree, V. & Robinson, M. (2001) Maturity, fecundity and reproductive potential of the European lobster (*Homarus gammarus*) in Ireland. *Journal of the Marine Biological Association of the UK* **81**, 61–68.

Ulrich, I., Muller, J., Schutt, C. & Buchholz, F. (2001) A study of population genetics in the European lobster, *Homarus gammarus* (Decapoda, Nephropidae). *Crustaceana* **74**, 825–837.

Van Wyngaarden, M., Snelgrove, P.V., DiBacco, C., Hamilton, L.C., Rodríguez-Ezpeleta, N., Jeffery, N.W., Stanley, R.R. & Bradbury, I.R. (2017) Identifying patterns of dispersal, connectivity and selection in the sea scallop, Placopecten magellanicus, using RADseq-derived SNPs. *Evolutionary Applications* **10**, 102–117.

Van Wyngaarden, M., Snelgrove, P.V., DiBacco, C., Hamilton, L.C., Rodríguez-Ezpeleta, N., Zhan, L., Beiko, R.G. & Bradbury, I.R. (2018) Oceanographic variation influences spatial genomic structure in the sea scallop, *Placopecten magellanicus*. *Ecology and Evolution* **8**, 2824–2841.

Vandamme, S.G., Griffiths, A.M., Taylor, S.A., Di Muri, C., Hankard, E.A., Towne, J.A., Watson, M. & Mariani, S. (2016) Sushi barcoding in the UK: another kettle of fish. *PeerJ* **4**, e1891.

Vendrami, D.L.J., Telesca, L., Weiss, M., Fawcett, K., Clark, M.S., Leese, F., Mcminn, C., Moore, H. & Hoffman, J.I. (2017) RAD sequencing resolves fine-scale population structure in a benthic invertebrate: implications for understanding phenotypic plasticity. *Royal Society Open Science* **4**, 160548.

Vieira, M.L.C., Santini, L., Diniz, A.L. & Munhoz, C.d.F. (2016) Microsatellite markers: what they mean and why they are so useful. *Genetics and Molecular Biology* **39**, 312–328.

Vikebø, F.B., Husebø, A.S., Slotte, A., Stenevik, E.K. & Lien, V.S. (2010) Effect of hatching date, vertical distribution, and interannual variation in physical forcing on northward displacement and temperature conditions of Norwegian spring-spawning herring larvae. *ICES Journal of Marine Science* **67**, 1948–1956.

Wahle, R.A. & Incze, L.S. (1997) Pre- and post-settlement processes in recruitment of the American lobster. *Journal of Experimental Marine Biology and Ecology* **217**, 179–207.

Wang, J., Santiago, E. & Caballero, A. (2016) Prediction and estimation of effective population size. *Heredity* **117**, 193–206.

Waples, R.S. & Gaggiotti, O. (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* **15**, 1419–1439.

Ward, R.D. (2006) The importance of identifying spatial population structure in restocking and stock enhancement programmes. *Fisheries Research* **80**, 9–18.

Watkinson, A.R. & Sutherland, W.J. (1995) Sources, sinks and psuedo-sinks. *Journal of Animal Ecology* **64**, 126–130.

Watson, H.V., McKeown, N.J., Coscia, I., Wootton, E. & Ironside, J.E. (2016) Population genetic structure of the European lobster (*Homarus gammarus*) in the Irish Sea and implications for the effectiveness of the first British marine protected area. *Fisheries Research* **183**, 287–293.

Watson, J.R., Mitarai, S., Siegel, D.A., Caselle, J.E., Dong, C. & McWilliams, J.C. (2010) Realized and potential larval connectivity in the southern California bight. *Marine Ecology Progress Series* **401**, 31–48.

Weersing, K. & Toonen, R. (2009) Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series* **393**, 1–12.

Weir, B.S. & Cockerham, C.C. (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.

Whitlock, M.C. & Lotterhos, K.E. (2015) Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of F_{ST}. *The American Naturalist* **186**, S24–S36.

Wielgoss, S., Taraschewski, H., Meyer, A. & Wirth, T. (2008) Population structure of the parasitic nematode *Anguillicola crassus*, an invader of declining North Atlantic eel stocks. *Molecular Ecology* **17**, 3478–3495.

Williams, J., Jenkins, G.P., Hindell, J.S. & Swearer, S.E. (2018) Fine-scale variability in elemental composition of estuarine water and otoliths: developing environmental markers for determining larval fish dispersal histories within estuaries. *Limnology and Oceanography* **63**, 262–277.

Williamson, D.H., Harrison, H.B., Almany, G.R., Berumen, M.L., Bode, M., Bonin, M.C., Choukroun, S., Doherty, P.J., Frisch, A.J., Saenz-Agudelo, P. & Jones, G.P. (2016) Large-scale, multidirectional larval connectivity among coral reef fish populations in the Great Barrier Reef Marine Park. *Molecular Ecology* **25**, 6039–6054.

Wilson, G.A. & Rannala, B. (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177–91.

Winter, D.J. (2012) MMOD: An R library for the calculation of population differentiation statistics. *Molecular Ecology Resources* **12**, 1158–1160.

Wolfram, K., Mark, F.C., John, U., Lucassen, M. & Pörtner, H.O. (2006) Microsatellite DNA variation indicates low levels of genetic differentiation among cuttlefish (*Sepia officinalis*) populations in the English Channel and the Bay of Biscay. *Comparative Biochemistry and Physiology, Part D* **1**, 375–383.

Wood, C. (2013) *Seasearch guide to sea anemones and corals of Britain and Ireland*. Wild Nature Press, Plymouth, United Kingdom.

Wright, D., Bishop, J.M., Matthee, C.A. & von der Heyden, S. (2015) Genetic isolation by distance reveals restricted dispersal across a range of life histories: implications for biodiversity conservation planning across highly variable marine environments. *Diversity and Distributions* **21**, 698–710.

Wright, S. (1932) The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress on Genetics* **1**, 356–366.

Wright, S. (1943) Isolation by distance. *Genetics* **28**, 114–138.

Xuereb, A., Benestan, L., Normandeau, E., Curtis, J.M., Bernatchez, L. & Fortin, M.J. (2018) Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). *Molecular Ecology* **27**, 2347–2364.

Young, H.S., Maxwell, S.M., Conners, M.G. & Shaffer, S.A. (2015) Pelagic marine protected areas protect foraging habitat for multiple breeding seabirds in the central Pacific. *Biological Conservation* **181**, 226–235.

Younger, J.L., Clucas, G.V., Kao, D., Rogers, A.D., Gharbi, K., Hart, T. & Miller, K.J. (2017) The challenges of detecting subtle population structure and its importance for the conservation of emperor penguins. *Molecular Ecology* **26**, 3883–3897.

Yu, H., You, X., Li, J., Zhang, X., Zhang, S., Jiang, S., Lin, X., Lin, H.R., Meng, Z. & Shi, Q. (2018) A genome-wide association study on growth traits in orangespotted grouper (*Epinephelus coioides*) with RAD-seq genotyping. *Science China Life Sciences* **61**, 934–946.